Malignant mesothelioma and asbestos-related lung cancer: diagnosis, prognosis and burden

Sjoukje van der Bij

Author: Sjoukje van der Bij
ISBN: 978-90-5335-618-0

Cover by: Anton Westbroek
Lay-out by: Nikki Vermeulen, Ridderprint BV, Ridderkerk, the Netherlands
Printed by: Ridderprint BV, Ridderkerk, the Netherlands

The studies described in this thesis were funded by the Dutch Institute for Asbestos Victims and the Department of Cardio-thoracic Surgery of the Academic Medical Center at the University of Amsterdam.

Financial support for the publication of this thesis by Maatschap van der Bij is gratefully acknowledged.

Malignant mesothelioma and asbestos-related lung cancer: diagnosis, prognosis and burden

Maligne mesothelioom en asbest-gerelateerde longkanker:
diagnose, prognose en omvang
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van derector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op donderdag 6 december 2012
des middags te 4.15 uur

door

Sjoukje van der Bij

geboren op 24 december 1978 te Arnhem

Promotoren: Prof. dr. K.G.M. Moons
Prof. dr. mr. dr. B.A.J.M. de Mol

Co-promotoren: Dr. ir. H. Koffijberg
Dr. ir. R.C.H. Vermeulen

Contents

Chapter 1

Introduction

Introduction

“It is now known that asbestos dust is one of the most dangerous dusts to which man is exposed” (The United States Bureau of Mines in 1932)

In the Western world, asbestos has been utilized by industrialized nations for over a century with a peak in the mid-1970s. It has been widely used in the building, shipping, and sound-proofing industries. In the 1950s and 1960s the hazardous effects of asbestos became widely recognized and accepted.[1;2] Since that time, it has taken many years to restrict the manufacturing and transport of asbestos. Not until the 1990s, the handling of asbestos dropped severely in the Western world due to directives on protecting workers exposed to asbestos. In 2005, a complete ban has been imposed in the whole European Union.[3] While the use of asbestos is now phased out in most Western countries, the use of asbestos in developing countries at present is even higher than historical use in Europe and North America in the 1970s.

To date, the hazards of asbestos in the Western world are not over yet as asbestos is virtually everywhere in the living environment. Asbestos may be released during accidents such as a fire or during reconstruction activities. Moreover, due to the regulations the asbestos industry has shifted from manufactory to removal work resulting in a small group of workers that still might be occupationally exposed. Potential exposure to asbestos may also cause panic among citizens. A recent example is the case in Utrecht, The Netherlands, where asbestos was found in and around an apartment block during renovation activities. About 150 people were ordered to leave their homes.[4]

In terms of health, the impact of the widespread use of asbestos in the Western world is still noticeable as asbestos is related to cancers that may occur many years after exposure. The most notably asbestos-related cancers are malignant mesothelioma and lung cancer. Malignant mesothelioma is almost always caused by asbestos exposure. The disease has a worse prognosis, and most affected individuals die within a year of diagnosis.[5-8] Mesothelioma mortality rates have been rising in Western countries over the past 20 years, and are expected to continue to rise till 2018.[9] In the Netherlands, the total death toll between 2000 and 2028 has been estimated to exceed 12,000 cases.[10] Individuals who have developed malignant mesothelioma as a result of asbestos exposure may be eligible to receive financial compensation for their losses and suffering by pursuing legal action. Many countries have a reimbursement system for patients with malignant mesothelioma. In the Netherlands, the Dutch Institute for Asbestos Victims was founded in 2000. Its primary task is to support malignant mesothelioma patients in the juridical claim process. To date, main challenges in the domain of malignant mesothelioma are to rapidly obtain a definite diagnosis, and to estimate and ideally increase the life expectancy of patients diagnosed with malignant mesothelioma. The first part of this thesis focuses on these two aspects. Although the

association between malignant mesothelioma and asbestos exposure is extremely strong, lung cancer is the most important asbestos-related cancer in terms of excess deaths. Ratios of 1:≥2 between malignant mesothelioma and excess lung cancers have been frequently observed in different cohorts.[11] However, the asbestos-related lung cancer burden is more difficult to quantify than for malignant mesothelioma due to the potential confounding role by other causes of lung cancer (notably smoking). The expected number of lung cancers due to asbestos exposure depends on the proportion of the population that is exposed to asbestos, the exposure levels and the associated lung cancer risk. As it remains hard to accurately assess the number of people exposed in the general population, the asbestos-related lung cancer risk is still uncertain, particularly in the lower exposure range. These issues are therefore further studied in this thesis.

Outline of this thesis

This thesis consists of several parts. The first part is related to the diagnostic and prognostic strategies of malignant mesothelioma. In **Chapter 2** we discuss the pathological and clinical elements comprising the diagnostic process of malignant mesothelioma patients currently implemented in the Netherlands, and their value in obtaining a definite diagnosis. To support diagnostic strategies for malignant mesothelioma we summarized the diagnostic accuracy of non-invasive markers for the diagnosis of malignant mesothelioma, in **Chapter 3**. In **Chapter 4** we identify factors related to the prognosis of malignant mesothelioma to optimize prognostic and reimbursement strategies.

The second part focuses on the asbestos-related lung cancer risk. In **Chapter 5** we estimate the exposure-response relationship between asbestos and lung cancer. Based on this relationship, and other relevant evidence, the burden of lung cancer was estimated with use of different methods which are presented in **Chapter 6.**

The third part of this thesis, **Chapter 7**, methodological considerations related to meta-analysis of diagnostic markers or tests are discussed, as we observed a large heterogeneity in the diagnostic accuracy of markers of malignant mesothelioma in chapter 3. We provide insight and guidance how to interpret, assess, and report the impact the impact of between-study variability in meta-analysis of diagnostic markers and tests.

The final **Chapter 8** provides concluding remarks and discusses implications for current and future policy making, practice and research.

Reference List

1. Doll R. Mortality from lung cancer in asbestos workers. Br J Ind Med 1955; 12:81-86.
2. Wagner JC, Sleggs CA, Marchand P. Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape Province. Br J Ind Med 1960; 17:260-271.
3. Virta RL. Worldwide asbestos supply and consumption trends from 1900 through 2003: U.S. Geological Survey Circular 1298. 2006.
4. Asbestos found in Utrecht homes. Dutch News; 23-7-2012. Retrieved from http://www.dutchnews.nl/news/archives/2012/07/asbestos_found_in_utrecht_home.php
5. Chapman A, Mulrennan S, Ladd B, Muers MF. Population based epidemiology and prognosis of mesothelioma in Leeds, UK. Thorax 2008; 63:435-439.
6. Mirabelli D, Roberti S, Gangemi M, Rosato R, Ricceri F, Merler E et al. Survival of peritoneal malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:194-200.
7. Montanaro F, Rosato R, Gangemi M, Roberti S, Ricceri F, Merler E et al. Survival of pleural malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:201-207.
8. Neumann V, Rutten A, Scharmach M, Muller KM, Fischer M. Factors influencing long-term survival in mesothelioma patients--results of the German mesothelioma register. Int Arch Occup Environ Health 2004; 77:191-199.
9. Peto J, Decarli A, La VC, Levi F, Negri E. The European mesothelioma epidemic. Br J Cancer 1999; 79:666-672.
10. Segura O, Burdorf A, Looman C. Update of predictions of mortality from pleural mesothelioma in the Netherlands. Occup Environ Med 2003; 60:50-55.
11. Henderson DW, Rodelsperger K, Woitowitz HJ, Leigh J. After Helsinki: a multidisciplinary review of the relationship between asbestos exposure and lung cancer, with emphasis on studies published during 1997-2004. Pathology 2004; 36:517-550.

Part 1

Diagnosis and prognosis of malignant mesothelioma

Chapter 2

Pathological and clinical assessment in the diagnosis of mesothelioma

S van der Bij
P Baas
MJ van de Vijver
BAJM de Mol
JA Burgers

Submitted

Abstract

Background: Apart of medical reasons, a definitive diagnosis of malignant mesothelioma may be required as a basis for a claim of financial compensation although a pathological source of conclusive evidence is missing. Clinical assessment of all available data is then the only option to come to a final conclusion. We evaluated the diagnostic work-up of a large cohort of Dutch patients who applied for financial compensation due to mesothelioma. We determined how often a pathological or clinical diagnosis can be made, and which factors are associated with making the final diagnosis malignant mesothelioma.

Methods: A flow diagram of the diagnostic work-up was constructed for patients that applied to the Dutch Institute for Asbestos Victims between 2005 and 2008 (N=1,498). Both pathological and clinical factors that may influence the diagnostic outcome were assessed.

Results: In 97 of the 1,498 patients (6%) no pathologic diagnosis could be established because of an uncertain diagnosis (N=54), inadequate (n=22) or unavailable tumor samples (N=21). A final pathological diagnosis of malignant mesothelioma could most often be made when biopsy samples were available compared to those in whom only cytological material was available. In patients in who no conclusive diagnosis could be made, clinical assessment was performed. Eighty percent of patients (66/83) who were clinically assessed were considered to have mesothelioma. None of the clinical features analyzed were strongly associated with a confirmed diagnosis of malignant mesothelioma.

Discussion: Our study shows that only in a small number of the patients who applied no pathologic diagnosis could be obtained. Based on judgment of clinical experts in the majority of these cases a near to certain diagnosis could be made. Moreover, it is reasonable to obtain biopsy material from patients to increase the chance to obtain a confirmed diagnosis. Therefore, it is important to refer patients early for diagnostic procedures.

Introduction

Malignant mesothelioma is a malignancy known for its long latency period after asbestos exposure.[1] Patients who have developed malignant mesothelioma as a result of occupational or environmental asbestos exposure may seek compensation for their losses and suffering by pursuing legal action. As the burden of malignant mesothelioma will remain high in the coming decade[2;3], compensation for those exposed in the past will remain an important issue.

The diagnosis of malignant mesothelioma is normally confirmed with use of pathologic material.[4] However, a definite pathological diagnosis may not always be feasible, either because of diagnostic difficulties or because of inadequate or unavailable tumor samples. To obtain a diagnosis, clinical assessment by some kind of 'diagnostic expert panel' is the only option to determine whether malignant mesothelioma is very likely or not.

In the Netherlands, patients with apparent malignant mesothelioma can apply to the Dutch Institute for Asbestos Victims for financial compensation. For each applicant, the diagnosis of malignant mesothelioma first needs to be confirmed by a 'national panel of pathologists', using both histological and cytological samples. If a diagnosis of malignant mesothelioma can not be made on the basis of cyological or histological evaluation (for whatever reason), subsequently a panel of 'clinical experts' evaluates all available clinical and radiological data, to ultimately determine whether the presence of malignant mesothelioma is more likely than some other diagnosis.[5] Accordingly, in The Netherlands, both patients with a pathologically or clinically confirmed diagnosis are candidates for a financial reimbursement. In this paper we evaluate the diagnostic work-up of a large cohort of almost 1,500 Dutch patients who applied for financial compensation, to determine how often a pathological or clinical diagnosis can be made, and which factors are associated with making a final diagnosis malignant mesothelioma.

Methods

Patients

The Dutch Institute for Asbestos Victims was founded in 2000 and its primary task is to support malignant mesothelioma patients in the legal claim process. Patients apply to the institute when they are diagnosed with, or are suspected, of a malignant mesothelioma based on judgment of the referring hospital. If this diagnosis is confirmed (see below), the patients or their relatives are entitled to financial compensation mediated by the Institute for Asbestos Victims. Since 2000 (until 2008), a total of 3,475 patients applied to the Institute for Asbestos Victims. Findings from the years 2000-2004 have been described before.[5] Here we focus on the diagnostic work-up and obtained final diagnoses, plus associated factors over the years 2005-2008 (N=1,498). From each patient, informed consent was obtained.

Diagnostic outcome

Pathological assessment

For each patient who applies for financial compensation, representative tumor samples are requested from the referring hospital, and reviewed by a national expert panel of pathologists: the so-called Dutch National Mesothelioma Panel (NMP). The reviewed slides may include cytological or histological (biopsy) material. The NMP classifies the diagnosis of each patient to one of the following categories:

I. Definite malignant mesothelioma
II. Probable malignant mesothelioma
III. Uncertain diagnosis of malignant mesothelioma; not able to differentiate malignant mesothelioma from e.g. mesothelial proliferation or an other type of malignancy
IV. No malignant mesothelioma (a diagnosis other than malignant mesothelioma)
V. Insufficient pathological material for making the diagnosis malignant mesothelioma

In The Netherlands, a patient can only be accepted for any financial compensation when the diagnosis malignant mesothelioma is confirmed (in case of category I and II). The request is rejected for all cases of category IV. In case of category III or V the clinical expert panel is subsequently asked to make a final diagnosis.

Clinical assessment

When pathological material is not available, insufficient or the pathological diagnosis by the NMP was uncertain a final diagnosis is reached by the so-called 'Mesothelioma Clinical Expert Panel of the Dutch Thoracic Society' (DTS).[5] This panel consists of 12-15 independent pulmonologists skilled in diagnosing malignant mesothelioma, who evaluate all available clinical and radiological data to conclude that either or not malignant mesothelioma is the (most likely) final diagnosis (yes/no diagnosis of malignant mesothelioma). Clinical features that are taken into account include e.g. gender, age, smoking status, asbestos exposure, chest pain, dyspnea, weight loss, progress of disease, other diseases that may explain symptoms. Radiological data may include features from X-thorax and CT-scans such as calcified pleural mass, irregular pleural thickening, interlobar fissure invasion, loss of volume of the hemithorax, pleural effusion. Finally, if available, pathological reports are considered.

Analyses

A flow diagram of the diagnostic work-up was constructed for the patients that applied to the Institute for Asbestos Victims between 2005 and 2008. Subsequently, both pathological and clinical factors that may influence the diagnostic outcome assessment were analyzed, using cross tabulations with Chi-square testing and Risk Ratio's with 95% Confidence Intervals.

Results

Patients

In the period between 2005 and 2008, 1,498 patients with apparent malignant mesothelioma applied for a financial compensation to the Dutch Institute for Asbestos Victims. After submission a diagnostic tract starts as shown in figure 1.

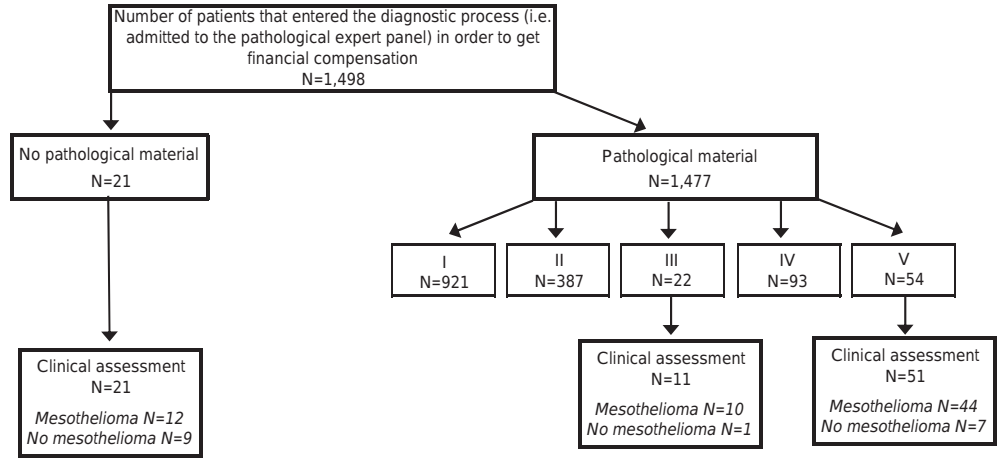


Figure 1: Flow diagram of the patients that entered the diagnostic process for getting financial compensation. Category I: definite malignant mesothelioma; Category II: probable malignant mesothelioma; Category III: uncertain diagnosis of malignant mesothelioma; Category IV: no malignant mesothelioma; Category V: Insufficient pathological material for making the diagnosis malignant mesothelioma.

Pathologic material was available of 1,477 patients. Among them, the NMP confirmed the diagnosis in 1,308 (89%) patients (category I and II of the flow diagram) and definitely ruled it out in 93 (6%) patients (category IV). The pathologic diagnosis remained uncertain in 76 patients because of diagnostic difficulties (category III (N=54)) or inadequate tumor samples (category V (N=22)). Moreover, no pathologic material was available for 21 patients. Thus, in 97 (6%) of the 1,498 patients no pathologic diagnosis could be established. Of these 97 patients, 83 patients underwent clinical assessment. A diagnosis of malignant mesothelioma based on clinical assessment was confirmed in 66 of these 83 patients (80%).

Most of the patients that underwent clinical assessment were alive at time of clinical assessment (66%). Reasons that patients did not have any pathologic material or only cytological material available were mainly due to a poor condition of the patient or unwillingness to undergo invasive diagnostic procedures. However, patients with pathologic material available but no established diagnosis of malignant mesothelioma had a higher probability to get a confirmed diagnosis of malignant mesothelioma based on clinical assessment compared to patients without any pathologic material available (54 of 62 patients (87%) versus 12 of 21 patients (57%)).

Factors influencing the diagnostic outcome

Association pathological material and final diagnosis in patients with pathological assessment

Table 1 and 2 show that patients with only cytological material available significantly more often did not score a category I or II diagnosis (malignant mesothelioma considered present) as compared to patients for whom (also) biopsy material was available. However, among patients with only cytological material available, a diagnosis could still be established in 74% (category I,II and IV). In 3% of patients that did have biopsy material available, no definite pathologic diagnosis could be reached. For almost all patients (85%) in whom the diagnosis was pathologically excluded, a diagnosis of carcinoma of the lung was made.

Table 1: Type of material used by the expert panel of pathologists in relation to the final diagnosis

	Cytological material available only	Histological material available only	Cytological plus histological material available	Autopsy material available	Total
	N=113	*N=1,276*	*N=75*	*N=13*	*N=1,477*
Pathologic diagnosis	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**
I. definite malignant mesothelioma	12 (11)	855 (67)	45 (60)	9 (69)	921 (62)
II. probable malignant mesothelioma	57 (50)	310 (24)	19 (25)	1 (8)	387 (26)
III. uncertain diagnosis of malignant mesothelioma	8 (7)	12 (1)	1 (1)	1 (8)	22 (1)
IV. no malignant mesothelioma (a diagnosis other than malignant mesothelioma)	15 (13)	70 (5)	7 (9)	1 (8)	93 (6)
V. Insufficient pathological material for making the diagnosis malignant mesothelioma	21 (19)	29 (2)	3 (4)	1 (8)	54 (4)

Table 2: Association between type of pathological material available and final diagnosis

	Only cytological material available *N=113*	Other (additional) material available *N=1,364 (1,276 +75+13)*	*Chi-square (p-value)*
Pathologic diagnosis	**N (%)**	**N (%)**	
Established diagnosis (category I, II and IV)	84 (74)	1,317 (97)	
No established diagnosis (category III and V)	29 (26)	47 (3)	
			105.5 (<0.001)

Association clinical features and final diagnosis in patients who underwent clinical assessment

Table 3 presents, among those patients that underwent clinical assessment, the relation between clinical features and the final diagnosis of malignant mesothelioma. To increase statistical power, we added also the 153 patients that underwent clinical assessment over the years 2000-2004 in this analysis.[5] Almost all features were significantly associated with a confirmed diagnosis of malignant mesothelioma (either present or absent), except for

age, gender, smoking status, asbestos and calcified pleural mass. However, none of these features were associated with a confirmed diagnosis to such an extent that this could solely determine the diagnosis. This suggests that the final diagnosis (ruling in our out malignant mesothelioma) based on clinical assessment is a multifactorial process.

Table 3: The relation between various patient and clinical features and achieving the diagnosis of malignant mesothelioma, among patients that underwent clinical assessment between 2000 and 2008 (N=238)[a]

Feature		Total	Diagnosis of malignant mesothelioma	Relation between clinical features and obtaining a diagnosis of malignant mesothelioma
		N	N (%)	Risk Ratio (95% CI)
Age	< 70 years	98	74 (76)	1.03 (0.88-1.19)
	≥ 70 years	140	103 (74)	
Smoker	yes or past	120	94 (78)	1.11 (0.96-1.29)
	no	118	83 (70)	
Gender	Male	233	174 (75)	1.24 (0.61-2.56)
	Female	5	3 (60)	
Asbestos	Yes	181	140 (77)	1.19 (0.97-1.46)
	No	57	37 (65)	
Chest pain	Yes	86	72 (84)	1.21 (1.05-1.40)*
	No	152	105 (69)	
Dyspneu	Yes	156	130 (83)	1.45 (1.19-1.77)*
	No	82	47 (57)	
Weight loss	Yes	84	69 (82)	1.17 (1.01-1.35)*
	No	154	108 (70)	
Disease progress corresponds with malignant mesothelioma	Yes	64	56 (88)	1.26 (1.10-1.44)*
	No	174	121 (70)	
No calcified pleural mass	Yes	149	116 (78)	1.14 (0.96-1.34)
	No	89	61 (69)	
Pleural effusion	Yes	172	142 (83)	1.56 (1.23-1.97)*
	No	66	35 (53)	
Irregular pleural thickening	Yes	190	150 (79)	1.40 (1.08-1.82)*
	No	48	27 (56)	
Interlobar fissure invasion	Yes	32	29 (91)	1.26 (1.10-1.45)*
	No	206	148 (72)	
contraction of the affected hemithorax	Yes	113	105 (93)	1.61 (1.38-1.89)*
	No	125	72 (58)	
No other diseases that can explain symptoms	Yes	212	163 (77)	1.43 (0.99-2.05)
	No	26	14 (54)	
Any suspicion of malignant mesothelioma based on pathology [b]	Yes	94	87 (93)	1.54 (1.21-1.97)*
	No	45	27 (60)	

[a] Includes all patients that underwent clinical assessment between 2000 and 2008 (i.e. an extra of 153 patients that underwent clinical assessment between 2000 and 2004 were added to the 83 patients with a clinical assessment between 2005 and 2008). [b] Patients without any pathologic report were not taken into account. *Significant at a p-value of 0.05.

Discussion

Our study shows that in about 6% of the patients with suspected malignant mesothelioma pathologic material was not available or insufficient for diagnosis. Using our diagnostic system a diagnosis of malignant mesothelioma with high probability could be made in 80% of these patients based on judgment of clinical experts that reviews all available clinical and radiological data.

Biopsy material is recommended as the reference standard in the diagnosis of malignant mesothelioma. We observed that patients with biopsy material were more likely to get a conclusive pathologic diagnosis compared to patients with only cytological material available. These results support the fact that cytological material might be sufficient for the diagnosis mesothelioma, but more often histology supplies superior material for a definite pathological conclusion. Therefore, biopsy material is preferred in the diagnosis of malignant mesothelioma.[4] Still 3% of patients of whom biopsy material was available did not get a definite pathologic diagnosis either. This suggests that the quantity and quality of biopsy material play an important role in the diagnosis of malignant mesothelioma.

The clinical manifestations of malignant mesothelioma are usually non-specific. Therefore, it is recommended not to use clinical assessment alone as diagnostic criteria.[4] However our study shows that in 6% of the patients no pathologic diagnosis could be obtained. Apart from medical decision making, a definitive diagnosis is important as a basis for a claim of financial compensation. Based on clinical consensus these patients had a high probability of malignant mesothelioma. Our results further show that none of the clinical features in isolation were highly related to a confirmed diagnosis. Hence, the diagnosis based on clinical features is a multivariable process. Studies should further elucidate the accuracy of the different combinations of clinical features in patients with suspected malignant mesothelioma. Also, for a fair clinical assessment, a sufficient number of assessors and regular quality control audits are recommended. Moreover, In this situation, it might be interesting to investigate the added value of serum markers beyond these clinical features in the diagnosis of malignant mesothelioma.[6]

Around the world, the Dutch system is quite unique in the fact that patients without confirmed pathologic diagnosis have the possibility to apply for a financial reimbursement based on a clinical diagnosis. For most countries a diagnosis based on clinical assessment is not routine (table 4). In France only a confirmation based on pathologic material is valid. In Japan, South Africa and Australia a confirmation by clinical features might be possible, however for only selected cases. Also, the identification of exposure to asbestos might be required to obtain compensation (table 4).
In conclusion, many countries have a financial system available for claims but patients are often only entitled for compensation if the diagnosis is pathologically confirmed. This may

result in a small part of patients that do not have the possibility to get any compensation during his or her life. Moreover, it is reasonable to obtain biopsy material to have the highest chance to get a confirmed diagnosis of malignant mesothelioma. Therefore, it is important to refer patients early for diagnostic procedures. If pathologic material is not sufficient, clinical assessment could be an option, especially if re-biopsy is not an option (anymore). Clinical assessment is a multivariable process and therefore should be carefully assessed by several experts in the field. It should be further elucidated how well clinical assessment by experts can indeed identify patients with malignant mesothelioma. This may help patients in whom no pathologic diagnosis can be obtained. In this way, all patients with malignant mesothelioma will get a fair process and the possibility to get compensation during his or her life.

Table 4: Reimbursement systems in different countries

	The Netherlands	Belgium	France	Japan	South Africa	Australia
reimbursement system	Institute for asbestos victims	Fund for occupationally related diseases; Asbestos fund	Fund for indemnisation of victims of asbestos (FIVA); Federation Francaise de Societes d'assureurs	Workers' accident compensation insurance; Asbestos-related health damage relief law	Statutory - Compensation for occupational Injuries and diseases act (COIDA) for non-miners; Statutory - Occupational diseases in mines and works act (ODMWA) for miners; Asbestos and Kgalagadi Relief Trusts	The commonwealth; State based systems
asbestos exposure required	no	yes (in practice however, all cases will be compensated by the Asbestos Fund, except those with proven exposure in countries other than Belgium)	no (in case of FIVA)	no (in case of asbestos related health damage relief law)	yes	yes
confirmation based on biopsy	yes	yes	yes	yes	yes	yes
confirmation based on cytology	yes	yes	very rarely	yes	exceptionally	occasionally
confirmation based on clinical features	yes, in 6% (see figure 1)	yes, when pathologic diagnosis is uncertain	no	exceptionally	exceptionally	exceptionally
review committee	pathology clinically	asbestos exposure, pathology, clinically	pathology	pathology	-	-
website(s)	www.asbestslachtoffers.nl	www.afa.fgov.be	www.fiva.fr; www.ffsa.fr	www.erca.go.jp/english/pdf/pamphlet_en.pdf	www.labour.gov.za/DOL; www.info.gov.za/acts/1993/a208-93.pdf www.asbestostrust.co.za/	www.ddb.nsw.org.au; www.asbestos.tas.gov.au/compensation; www.comcare.gov.au; wwwsafeworkaustralia.gov.au

Acknowledgements

This study was funded by the Institute for Asbestos Victims, the Netherlands.

We thank Prof. Dr. N. van Zandwijk, Prof Dr. M. Praet, Prof. Dr. J. van Meerbeeck, Prof. Dr. A Scherpereel, Ms. H. Mauss of the FIVA, Dr. T. Nakano, Mr. S Furuya, Dr. J Te Water Naude, Dr. M. Vandeweerdt for supplying information on country specific guidelines for financial reimbursement for patients with malignant mesothelioma.

Conflict of interest

B.A.J.M.M is a board member of the Institute for Asbestos Victims. M.J.V is chairman of the Dutch National Mesothelioma Panel. J.A.B. is chairman of the Mesothelioma Group of the Dutch Thoracic Society.

Reference List

1. Marinaccio A, Binazzi A, Cauzillo G, Cavone D, Zotti RD, Ferrante P et al. Analysis of latency time and its determinants in asbestos related malignant mesothelioma cases of the Italian register. Eur J Cancer 2007; 43:2722-2728.
2. Peto J, Decarli A, La VC, Levi F, Negri E. The European mesothelioma epidemic. Br J Cancer 1999; 79:666-672.
3. Segura O, Burdorf A, Looman C. Update of predictions of mortality from pleural mesothelioma in the Netherlands. Occup Environ Med 2003; 60:50-55.
4. Scherpereel A, Astoul P, Baas P, Berghmans T, Clayson H, de Vuyst P et al. Guidelines of the European Respiratory Society and the European Society of Thoracic Surgeons for the management of malignant pleural mesothelioma. Eur Respir J 2010; 35:479-495.
5. Baas P, van 't Hullenaar N, Wagenaar J, Kaajan JP, Koolen M, Schrijver M et al. Occupational asbestos exposure: how to deal with suspected mesothelioma cases--the Dutch approach. Ann Oncol 2006; 17:848-852.
6. van der Bij S, Schaake E, Koffijberg H, Burgers JA, de Mol BA, Moons KG. Markers for the non-invasive diagnosis of mesothelioma: a systematic review. Br J Cancer 2011; 104:1325-1333.

Chapter 3

Markers for the non-invasive diagnosis of mesothelioma: a systematic review

S van der Bij
E Schaake
H Koffijberg
JA Burgers
BAJM de Mol
KGM Moons

Published in Br J Cancer 2011: 104(8): 1325-33

Abstract

Background: Numerous markers have been evaluated to facilitate the non-invasive diagnostic work-up of mesothelioma. The purpose of this study was to conduct a structured review of the diagnostic performance of non-invasive marker tests for the detection of mesothelioma in patients with suspected mesothelioma.

Methods: Studies on the diagnostic accuracy of serum and cytological markers published till 31 December 2009 available in either Pubmed or Embase to detect or exclude the presence of mesothelioma were extracted. Study quality was assessed with use of the QUADAS criteria.

Results: 82 articles were included in this systemic review. Overall, quality of the incorporated studies to address our objective was poor. The most frequently studied immunohistochemical markers for cytological analysis were EMA, Ber-Ep4, CEA, and calretinin. The most frequently investigated serum marker was SMRP. CEA, Ber-EP4 and calretinin were most valuable in discriminating mesothelioma from other malignant diseases. EMA and SMRP were most valuable in discriminating mesothelioma from non-malignant diseases. No marker performed well in discriminating between mesothelioma and all other diseases.

Conclusion: Currently, there is only limited evidence to properly assess the value of non-invasive marker tests in the diagnosis of mesothelioma. Studies were of limited value to address our objective and results showed considerable unexplained study heterogeneity.

Introduction

The diagnosis of mesothelioma is not straightforward. The symptoms are non-specific, and only in experienced centers pleural fluid cytology is a reliable diagnostic tool. Hence, most patients ultimately require invasive procedures such as core-needle or open biopsy, or video assisted thoracoscopy to facilitate histological examination as 'gold' standard for diagnosis.[1-3] However, a biopsy may complicate subsequent disease management by seeding tumor cells or may be unfeasible due to poor condition of the patient. Therefore, it would be valuable to have non-invasive diagnostic procedures that accurately confirm or exclude the diagnosis of mesothelioma.

Accordingly, innumerable non-invasive markers have emerged, based on the increasing understanding of the molecular and biological pathways of mesothelioma, and studied in numerous studies. These include many immunohistochemical markers that have been tested for their property to establish the diagnosis of mesothelioma on cytological grounds.[4] Promising other tests are genetic markers and serum markers such as soluble mesothelin-related protein (SMRP) and megakaryocyte potentiating factor (MPF).[5-8]

However, estimated diagnostic accuracy of identical markers varies widely between studies. Therefore it remains unclear which marker has a superior performance. Nevertheless, several markers have already entered the market and are used in clinical practice. In contrast, others disappeared after initial promising results. As a result, current diagnostic strategies for mesothelioma involving markers are likely to be suboptimal. Therefore, we conducted a systematic review to summarize the literature on the diagnostic accuracy of serum and cytological markers for the diagnosis of mesothelioma.

Methods

Search strategy

The systematic search addressed articles with information on markers in serum and effusions to include or exclude the presence of mesothelioma published till 31 December 2009. The search was carried out with Medline and Pubmed (see appendix 1 for search strategy).Duplicates from Medline and Embase were deleted automatically and manually with Reference Manager v11 (Thomson Reuters, New York, USA).

Markers (index tests)

To facilitate the analysis, and to allow a more appropriate comparison between the studies, we divided the non-invasive markers into four groups: serum markers; effusion markers, i.e. pleural and peritoneal fluid markers; immunohistochemical markers used for cytological analysis of effusion samples; and genetic markers.

Selection

To be eligible for inclusion, studies had to fulfil all of the following criteria:

1. The study should be an original report in English (i.e. letters, editorials, case-reports, tutorials, reviews and non-English studies were excluded);
2. The study should assess the ability of one or more markers to detect or exclude the presence of mesothelioma, and only involving non-invasive marker tests. Studies in which marker tests were assessed in tissue biopsies, pelvic washings or more than 10% fine needle aspirates (FNAs) were not included;
3. The diagnosis of mesothelioma had to be confirmed on at least cytology and/or histology.
4. The study should have a minimal sample size of 10 mesothelioma patients;
5. The study should provide sufficient data to (re)construct a two-by-two contingency table to estimate the marker's diagnostic accuracy.

Studies reporting ≥10% more specimens than study patients indicating that more than one specimen per patient was used, were excluded. Furthermore, studies investigating markers for purpose of screening or surveillance of high risk study populations were excluded.

The article selection was performed in two consecutive phases: title and abstract assessment (one reviewer, S.B) and full article assessment (two independent reviewers, S.B and E.S.).

Data extraction

If a study was included, the two reviewers independently extracted the following elements from the article: overall study characteristics, e.g. author(s), institution, date of publication, recruitment setting, study design and study years; participant characteristics, e.g. description of the mesothelioma patients and comparison group; details of the index marker test including the positive versus negative cut-off value; type of reference test used to confirm the presence or absence of mesothelioma.

The number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) were extracted and used to construct a two-by-two table, if possible separately for each comparison group. Comparison groups were summarized to either other malignancies or no malignancies, which could include also healthy participants. If more than one cut-off value was used, we selected the value closest to the cut-off corresponding with 95% specificity (avoiding false positives as much as possible). For immunohistochemical markers, we selected the value closest to the 10% cut-off according to the percentage of cells exhibiting staining (as it is a frequently used value in immunocytology and implies that samples were considered positive for the marker if at least 10% of malignant mesothelioma cells were immunohistochemically stained). Data of the diagnostic value of a combination of markers were not extracted.

Discrepancies between the two reviewers were resolved by consensus. If needed a third and fourth reviewer (H.K., K.G.M.M) resolved the remaining discrepancies. When studies with overlapping data sets were published, preference was given to those studies which had

the highest number of mesothelioma patients or used malignancy as a comparison group (which better reflects clinical practice). If a study evaluated various markers and results of a subset of these markers were published in a more recent study, then only the results of the duplicate markers were excluded from the first study.

Quality assessment

The methodological quality of each included study was independently assessed by the two reviewers using the QUADAS instrument (see appendix 2), a widely accepted and validated tool for the quality assessment of diagnostic accuracy studies in systematic reviews.[9] In case of doubt, a third or fourth reviewer was consulted (H.K. and K.G.M.M).

Data synthesis

Results were summarized per type of marker and per comparison group (i.e. other malignancies or no malignancies). Markers reported in at least 6 studies were described more comprehensively. As is common in diagnostic systematic reviews and meta-analysis, we used sensitivity and specificity as our primary measures of association. Sensitivity was calculated by dividing TP by (TP+FN) and specificity by dividing TN by (FP+TN) from the (re)constructed 2 by 2 tables. Associated 95% confidence intervals (CIs) were assessed using the Wilson score method.[10] To graphically present the results, estimates of sensitivity and specificity of a single marker across studies were summarized in a receiver operating characteristic (ROC) graph, plotting the markers' sensitivity on the y-axis against the 1-specificity on the x-axis. When different studies on the same biomarker show different sensitivity and specificity, it does not necessarily mean that the results are different or heterogeneous; they might simply have used a different (explicit or implicit) cut-off value for marker-positivity. As with a change in cut-off value the sensitivity and specificity commonly increase or decrease in opposite directions (negative correlation), the ROC curve for such marker should show a concave, shoulder like pattern. For each marker with different sensitivities and specificities plotted in ROC space, we quantify whether this could be explained by such threshold effect by estimating the (negative) correlation between sensitivity and specificity. This was done on the logit scale using the bivariate model.[11] All analyses were performed in SAS statistical packages, version 9.1 (SAS Institute Inc., Cary, NC, USA).

Results

Search results

Our search yielded 1642 hits, of which 307 were eligible for inclusion based on title and abstract. After assessment of the full text articles, 224 articles were discarded for various reasons (see figure 1). Thus, this review included 82 articles (see appendix 3): 36 articles that evaluated serum or effusion markers, 41 on immunohistochemical markers, 2 studies on genetic markers and 3 studies on different types of markers. Most immunohistochemical studies included epitheloid and biphasic mesotheliomas.

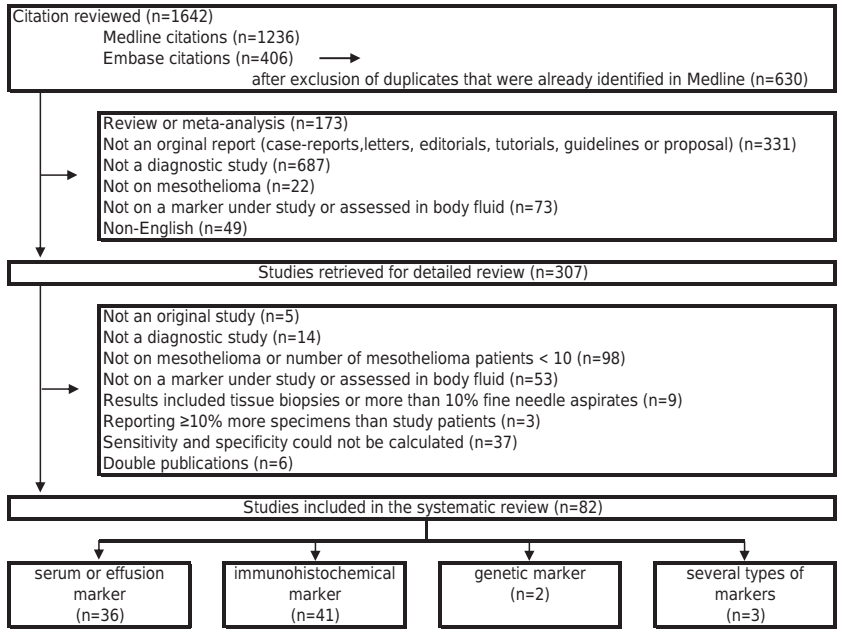


Figure 1: Flowchart of the selection of the relevant articles
Serum and effusion markers include tests to detect serum and effusion marker levels; immunohistochemical markers include marker tests used for cytological analysis of effusion samples; genetic markers include polymerase chain reaction tests to detect specific gene expressions or fluorescence in situ hybridization (FISH) tests to detect gene deletions with the use of specialized gene probes.

Study quality

The methodological quality of the studies with focus on the objective of this review was generally poor and is shown in figure 2, with specific details in table 1 (references to these studies are prefaced by an 'r' and listed in appendix 3). Only three articles were identified that adequately selected a representative cohort of consecutive patients suspected for mesothelioma.[r3;r22;r76] Of these, two articles were based on one prospective French study.[r3;r22] Other studies used a case-control design (n=70), or a cohort of patients with pleural effusions (n=9). Due to these designs nearly all studies (88%) suffered from the well described and problematic disease verification bias.[12-17] Furthermore, most studies did not have an adequate description of the patient selection procedure, characteristics of the study participants, the reference standard and the used cut-off value of the marker. The time between index test (marker) and reference test, as well as the availability of other clinical data (as is commonly encountered in practice) were also poorly reported. Blinding for the results of the marker (index test) when interpreting the reference test (and vice versa) was fulfilled in about 55% of the studies.

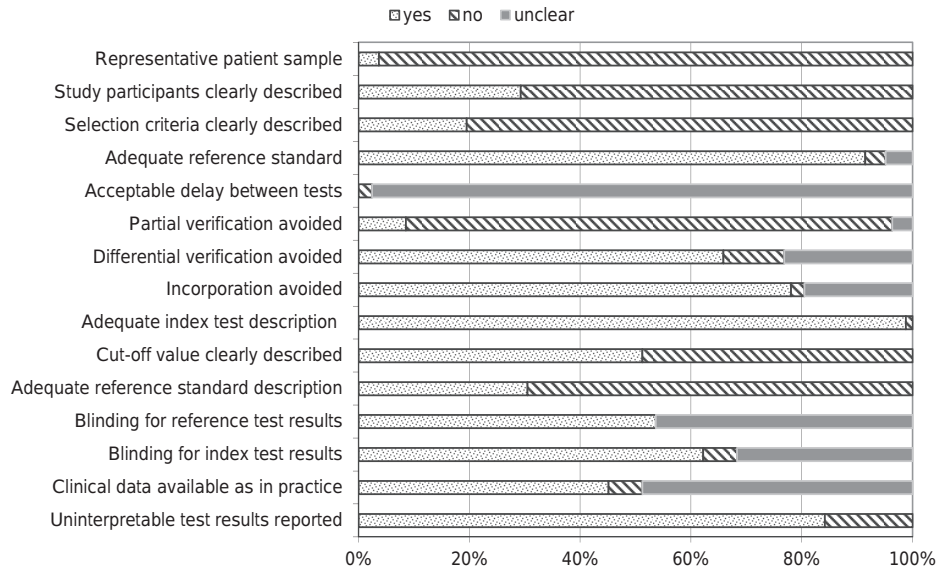


Figure 2: Summary of quality of the included studies according to the QUADAS criteria (see Appendix)

Table 1: Study characteristics and quality of included studies (ordered by year of study)

1st Author-year	Study design	Index test	Quality assessment^												
			1a	1b	2	3	4	5	6	7	8a	8b	9	10	11	12	13
Studying serum or effusion markers																	
Aleman - 2009[r1]	case-control	SMRP[(e)]	2	2	1	1	3	2	1	1	1	1	2	1	1	1	1
Davies - 2009[r2]	prospective cohort of consecutive patients with pleural effusion, suspected of pleural malignancy	SMRP[(e)]	2	1	1	1	3	2	1	1	1	1	2	1	1	1	1
Grigoriu - 2009[r3 #]	cohort of patients with suspected mesothelioma	HA[(es)]	1	1	1	1	3	1	1	1	1	2	1	1	3	1	2
Rodriquez Portal - 2009[r4]	prospective case-control	SMRP[(s)]	2	1	1	1	3	2	1	1	1	1	2	1	1	1	1
Shigematsu - 2009[r5]	case-control	Gene-X[(s)], THBS-2[(s)]	2	2	2	1	3	2	1	1	1	1	2	1	1	1	1
Amati - 2008[r6]	prospective case-control	80HdG[(s)], HGF[(s)], PDGFß[(s)], SMRP[(s)], VEGFß[(s)], bFGF[(s)]	2	1	1	1	3	2	1	1	1	2	2	1	1	1	1
Creaney - 2008[r7]	case-control	MPF[(s)], SMRP[(s)], osteopontin[(s)]	2	1	2	1	3	2	1	1	1	2	2	1	1	1	2
Iwahori - 2008[r8]	case-control	MPF[(s)], SMRP[(s)]	2	2	2	1	3	2	1	1	1	1	2	1	1	1	1
Pass - 2008[r9]	case-control	SMRP[(es)]	2	1	2	1	3	2	1	1	1	1	2	1	1	1	2
Schneider - 2008[r10]	prospective case-control	SMRP[(s)]	2	1	2	1	3	2	3	1	1	1	2	1	3	1	1
Creaney - 2007[r11 #]	case-control	CA125[(s)]	2	2	2	1	2	2	1	1	1	1	1	1	1	1	1
Creaney - 2007[r12 $]	retrospective cohort of consecutive patients with pleural effusion	SMRP[(e)]	2	1	1	1	3	1	1	1	1	1	2	1	1	1	2

Table 1: Study characteristics and quality of included studies (ordered by year of study) *(Continued)*

1st Author-year	Study design	Index test	Quality assessment^														
			1a	1b	2	3	4	5	6	7	8a	8b	9	10	11	12	13
Cristaudo - 2007[r13]	case-control	SMRP[(s)]	2	1	2	1	3	2	1	1	1	1	2	1	1	1	1
Di Serio - 2007[r14]	case-control	SMRP[(s)]	2	1	2	1	3	2	3	1	1	1	2	1	3	1	1
Grigoriu - 2007[r15 #]	case-control	osteopontin[(s)]	2	1	1	1	3	1	1	1	1	2	2	1	3	1	2
Shiomi - 2007[r16]	prospective case-control	MPF[(s)]	2	1	1	1	3	2	1	1	1	1	1	1	1	1	1
Van den Heuvel - 2007[r17]	retrospective case-control	CEA[(s)], CYFRA21-1[(s)], SMRP[(s)]	2	1	1	1	3	2	1	1	1	1	2	1	1	1	1
Welker - 2007[r18]	case-control	HA[(e)]	2	2	2	2	3	2	1	1	1	1	1	1	3	1	1
Onda - 2006[r19]	retrospective case-control	MPF[(s)]	2	1	2	1	3	2	1	1	1	1	2	1	1	1	1
Filiberti - 2005[r20]	prospective case-control	PDGF-AB[(s)]	2	1	2	1	3	2	1	1	1	1	1	1	1	1	1
Pass - 2005[r21]	case-control	osteopontin[(s)]	2	1	1	1	3	2	1	1	1	1	1	1	1	1	1
Scherpereel - 2005[r22]	prospective cohort of consecutive patients with suspected or recently diagnosed mesothelioma	SMRP[(es)]	1	1	1	1	3	1	1	1	1	1	2	1	3	1	1
Neri - 2003[r23]	prospective case-control	p53[(s)]	2	1	1	1	3	1	1	1	1	1	2	1	1	1	1
Villena - 2003[r24]	prospective cohort of patients with pleural effusion	CA15-3[(e)], CA549[(e)], CA72-4[(e)], CEA[(e)]	2	1	1	1	3	3	1	1	1	1	1	1	1	1	1
Creaney - 2001[r25]	case-control	p53[(s)]	2	2	2	1	3	2	1	1	1	1	2	1	1	1	1
Paganuzzi - 2001[r26]	cohort of consecutive patients with pleural effusion	CEA[(e)], CYFRA21-1[(e)]	2	2	2	1	3	1	3	1	1	1	2	1	3	1	1
Fuhrman - 2000[r27]	prospective case-control	CEA[(es)], HA[(e)]	2	2	2	1	3	2	1	1	1	1	2	1	1	1	2
Alatas - 1999[r28]	case-control	CA15-3[(es)], CA19-9[(e)], CEA[(es)], CYFRA21-1[(es)], NSE[(es)], TSA[(es)]	2	1	2	1	3	2	1	1	1	1	2	1	1	1	1
Miedouge - 1999[r29]	retrospective case-control	CA15-3[(e)], CA19-9[(e)], CA72-4[(e)], CEA[(e)], CYFRA21-1[(e)], NSE[(e)], SCC[(e)]	2	2	1	1	3	2	3	1	1	1	1	1	1	1	1
Nisman - 1998[r30]	case-control	CEA[(s)], CYFRA21-1[(s)], TPS[(s)]	2	2	2	1	3	2	1	1	1	1	1	1	1	1	1
Atagi - 1997[r31]	prospective cohort of consecutive patients with pleural effusion or previously diagnosed mesothelioma	CEA[(e)], HA[(e)]	2	2	1	1	3	1	3	3	1	1	1	1	3	1	2
Ebert - 1997[r32]	prospective case-control	CEA[(s)], CYFRA21-1[(s)], NSE[(s)], TPA-M[(s)], TPS[(s)]	2	2	2	1	3	2	1	1	1	1	1	1	1	1	2
Shijubo - 1995[r33]	case-control	CEA[(e)], SP-A[(e)]	2	2	2	1	3	2	1	1	1	1	2	1	1	1	1
Villena - 1995[r34]	prospective cohort of patients with pleural effusion	CA15-3[(e)], CA19-9[(e)], CA72-4[(e)], CEA[(e)]	2	1	1	1	3	3	1	1	1	1	2	1	1	1	1
Whitaker - 1986[r35]	retrospective case-control	CEA[(e)]	2	2	2	1	3	2	1	1	1	1	1	1	1	1	1

Table 1: Study characteristics and quality of included studies (ordered by year of study) *(Continued)*

1st Author-year	Study design	Index test	Quality assessment^														
			1a	1b	2	3	4	5	6	7	8a	8b	9	10	11	12	13
Fravelli - 1984[r36]	cohort of patients with pleural effusion	CEA[(e)]	2	1	2	1	3	3	1	1	1	1	1	1	1	1	1
Studying immunohistochemical markers																	
Shen - 2009[r37]	retrospective case-control	EMA, Glut-1m, Glut-1p, XIAP	2	2	2	1	3	2	3	3	1	2	2	1	1	3	1
Slipicevic - 2009[r38]	case-control	IGF-II, IGFBP3	2	2	2	1	3	2	3	3	1	2	2	3	3	3	1
Yuan - 2009[r39]	case-control	B72-3, Ber-EP4, EMA, Tenascin-X, calretinin	2	2	2	1	3	2	1	1	1	2	2	3	1	3	1
Bhalla - 2007[r40]	retrospective case-control	CK5, D2-40, calretinin, podoplanin	2	2	2	1	3	2	1	1	1	2	2	3	1	3	1
Facchetti - 2007[r41]	retrospective cohort of patients with effusion	claudin4	2	2	2	1	3	2	1	1	1	1	1	3	1	3	1
Grefte - 2007[r42]	retrospective case-control	B72-3, Ber-EP4, CEA, EMA, HMFG-2, calretinin	2	2	2	1	3	2	3	3	1	2	2	1	3	2	1
Kleinberg - 2007[r43]	retrospective case-control	claudin1, claudin3	2	2	2	1	3	2	1	1	1	2	2	3	1	3	1
Pu - 2007[r44]	retrospective case-control	MOC-31, WT-1, mesothelin, p63	2	2	2	1	3	2	1	1	1	2	2	1	1	3	2
Shield - 2007[r45]	retrospective case-control	CK5/6, calretinin	2	2	2	1	3	2	2	3	1	2	1	3	2	3	1
Aerts - 2006[r46]	prospective case-control	Ber-EP4, CEA, EMA, TAG-72	2	2	2	1	3	2	3	3	1	2	2	1	3	2	1
Bassarova - 2006[r47]	case-control	D2-40	2	2	2	1	3	2	3	3	1	2	2	3	3	3	1
Li - 2006[r48]	retrospective case-control	Ber-EP4, CAM5-2, CEA, CK5/6, K903, calretinin	2	2	2	1	3	2	3	3	1	2	2	3	3	3	1
Saad - 2006[r49]	retrospective case-control	CK5/6, D2-40, TTF-1, WT-1, calretinin, p63	2	2	2	1	3	2	3	1	1	2	1	1	3	2	1
Sivertsen - 2006[r50]	case-control	E-cadherin, N-cadherin, P-cadherin	2	2	2	1	3	2	3	3	2	2	2	3	3	3	1
Afify - 2005[r51]	retrospective case-control	CD44S, HA	2	2	2	1	3	2	1	1	1	2	1	3	1	3	1
Hecht - 2005[r52]	retrospective case-control	MOC-31	2	2	2	2	3	2	1	1	1	2	2	3	1	3	1
Saad - 2005[r53]	retrospective case-control	EMA	2	2	2	1	3	2	1	1	1	1	2	1	1	2	1
Saqi - 2005[r54]	case-control	CD138	2	2	2	1	3	2	1	1	1	1	2	3	1	3	1
Schonherr - 2004[r55]	case-control	Ki67	2	2	2	1	3	2	1	1	1	1	1	3	1	3	1
Afify - 2002[r56]	retrospective case-control	TTF-1	2	2	2	1	3	2	1	1	1	2	1	3	1	3	1
Afify - 2002[r57]	retrospective case-control	actin, desmin, myogenin, myoglobin	2	2	2	1	3	2	1	1	1	2	1	3	1	3	1

[(e)] assessed in effusion; [(es)] assessed in effusion and serum; [(es)] assessed in effusion and serum; [#] also studied other markers that we did not incorporate due to overlap with other studies; [$] also studied SMRP in serum that we did not incorporated due to overlap with other studies ^see appendix for criteria on quality assessment, items were scored 1=yes, 2=no, 3=unclear.

Investigated markers

Appendix 4 provide a complete summary of the performance of all markers, across the included studies. In total 54 immunohistochemical markers, 21 serum markers, 12 effusion markers and one genetic marker were identified. The most frequently evaluated immunohistochemical marker was EMA followed by BER-EP4, CEA, and calretinin (appendix 4 table 2.3). Among serum markers, the most frequently investigated were SMRP and CEA, (appendix table 2.1) and among effusion markers CEA, CA15-3, HA and SMRP (appendix 4 table 2.2). Results on genetic markers were sparse (appendix 4 table 2.4). The number of eligible papers allowed a closer evaluation of SMRP in serum and CEA in effusion as well as the immunohistochemical value of CEA, Ber-EP4, calretinin, and EMA.

Figures 3 and 4 show the ROC space plots for the SMRP in serum and CEA in effusions, and the immunohistochemical markers Ber-Ep4, CEA, EMA and calretinin. In figure 3 their performance to discriminate mesothelioma from other malignant diseases is shown, and in figure 4 the performance to discriminate mesothelioma from non-malignancies. From these figures a clear threshold effect seems apparent for SMRP, meaning that the variation between studies is probably due to differences in the applied positivity threshold.

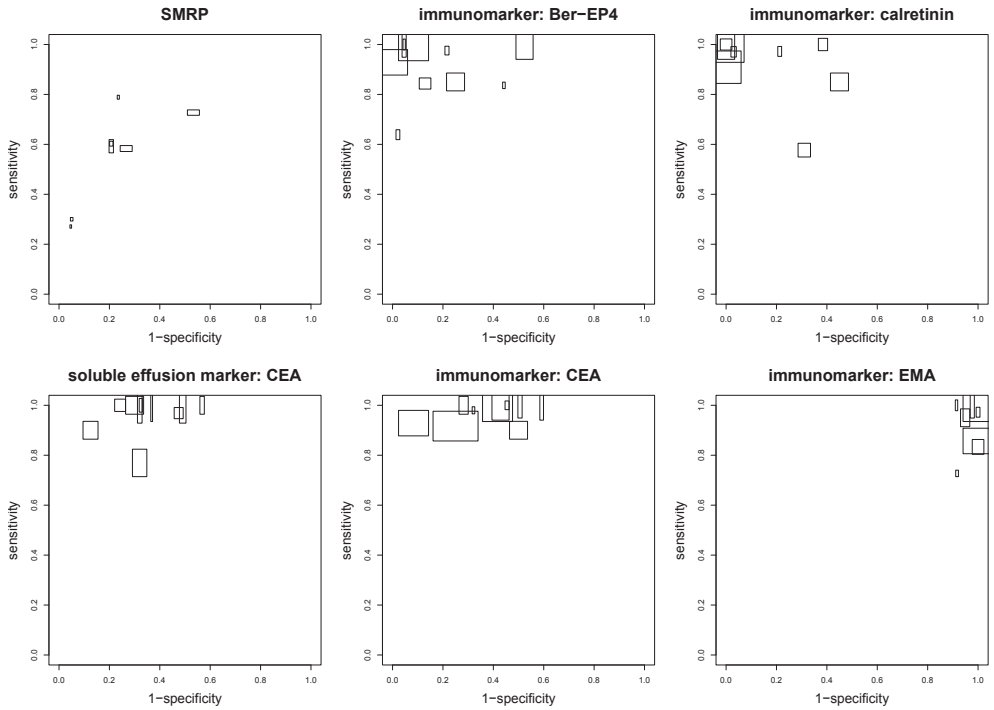


Figure 3: Sensitivity against 1-specificity in ROC space to discriminate mesothelioma from other malignant diseases. The height of the blocks is proportional to the reciprocal of the number of mesothelioma patients (mesothelioma yes subjects) and the width of the blocks is proportional to the reciprocal of the number of patients with other malignant diseases (mesothelioma no subjects).

Studies with a higher threshold mostly produced higher sensitivities and lower specificities. This finding is supported by the significant negative correlations between the logit sensitivity and logit specificity (-0.95, 95% CI:-0.99 - -0.27 in figure 3 and -1.00, 95% CI:-1.00 - -0.99 in figure 4). Overall, SMRP levels were lower among sarcomatoid mesothelioma compared to the other types (data not shown).

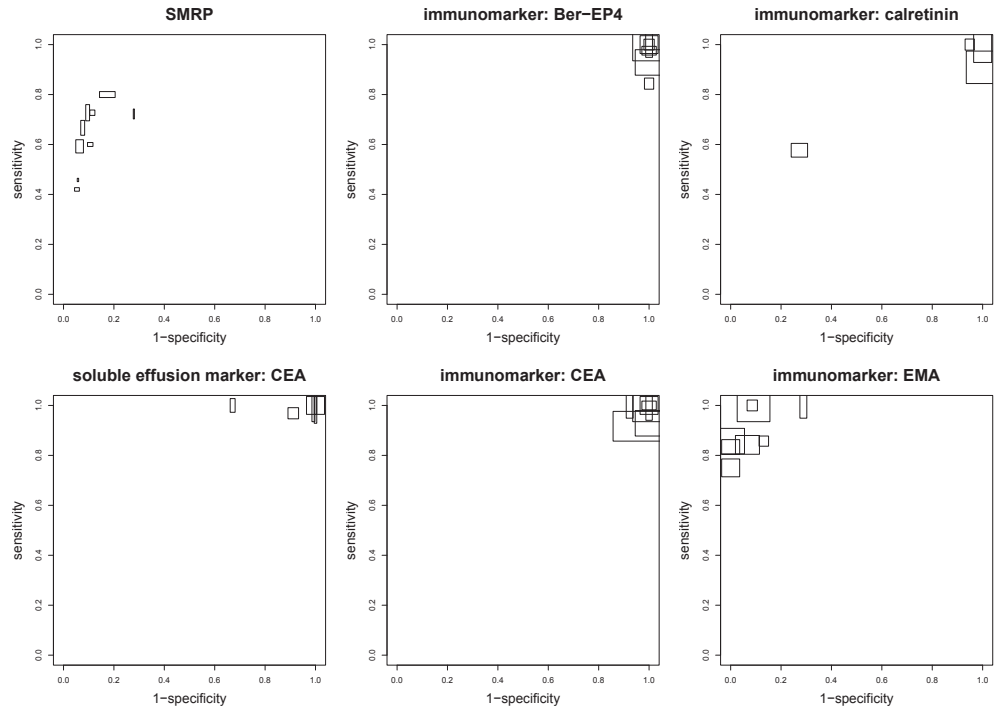


Figure 4: Sensitivity against 1-specificity in ROC space to discriminate mesothelioma from non-malignancy. The height of the blocks is proportional to the reciprocal of the number of mesothelioma patients (mesothelioma yes subjects) and the width of the blocks is proportional to the reciprocal of the number of non-malignant patients (mesothelioma no subjects).

In all CEA studies, effusion CEA levels lower than 40 ng/ml were compatible with both non-malignancy and mesothelioma. Discrimination between mesothelioma and non-malignancy based on CEA levels was therefore poor (figure 4). CEA levels among other malignancies were in general higher than in mesothelioma patients. Figure 3 shows that the specificity of CEA (i.e. the proportion of patients with other malignant diseases above a specific cut-off point) varied widely among studies and ranged from 43% (95% CI: 33-54) to 88% (95% CI: 69-96). These differences could only partly be explained by differences in the applied cut-off value (correlation was not significant), and by the type of other malignancies included in the control group.

The immunohistochemical markers Ber-EP4, CEA and calretinin can be useful in discriminating mesothelioma from other malignant diseases (figure 3), while EMA can be useful in discriminating mesothelioma from non-malignant diseases (figure 4). Specificity of Ber-Ep4 and CEA was more heterogeneous than sensitivity and sensitivity was, in general, high (figure 3). For calretinin the sensitivity ranged from 85% to 100% except for the study of Simsir et al. In that study [r65] calretinin staining was much lower among mesothelioma and benign samples.

EMA had a positive cytoplasmic or membranous staining in the majority of the papers, ranging from 73% to 100% among mesothelioma patients and from 91% to 100% among other malignant diseases. Four studies[r39;r42;r75; r82] made a distinction in staining pattern as well, showing that a membranous staining EMA pattern was mainly observed in mesothelioma patients (55%-92%) and not in other malignant diseases (<20%) (appendix 4 table 2.3). Discriminating mesothelioma from non-malignant diseases based on EMA yielded high sensitivity and specificity (figure 4). For EMA the correlation between logit sensitivity and specificity was non-significant in figure 4 (-0.56 (95% CI: -0.92-0.30)).

Direct marker comparisons

Some of the studies evaluated multiple markers on the same patients. Two studies[r42;r48] evaluated both the accuracy of calretinin and CEA. To discriminate mesothelioma from other malignant diseases both studies showed that specificity was higher for calretinin (in both studies: 100%) compared to CEA (in both about 58%). Corresponding sensitivities were 91% and 100% for calretinin and 100% (in both studies) for CEA.

Three studies[r39;r48;r80] in which calretinin and Ber-Ep4 were assessed showed that calretinin was a better discriminator than Ber-EP4, whereas one other study[r42] showed a similar performance of both markers.

Seven other studies directly compared the immunohistochemical markers CEA and Ber-EP4,[;r46;r48;r64;r67;r73;r82] Sensitivity values were highest for CEA, and in five of the seven studies [r42;r46;r64;r67;r73] Ber-Ep4 yielded the highest specificity.

No robust conclusion could be drawn on the relative performance of markers across comparative studies, due to large differences in study methods and heterogeneity of the results (table 1, appendix 4 table 2.3).

Discussion

We systematically reviewed all available evidence on the diagnostic performance of markers in serum, pleural fluid and ascites, used to non-invasively discriminate mesothelioma from non-mesothelioma disorders. Numerous markers have been assessed. SMRP, CEA, Ber-EP4, calretinin and EMA were studied most frequently. We found that the majority of studies had an exploratory design and as such showed a rather poor reporting and low quality as scored

by the QUADAS instrument for assessing methodological quality of individual studies in diagnostic reviews. Nevertheless, despite this, our analyses indicate that the most valuable markers appear to be CEA, Ber-EP4 and calretinin to discriminate mesothelioma from other malignant diseases. EMA and SMRP were most valuable in discriminating mesothelioma from non-malignant diseases. None of the markers performed well to differentiate mesothelioma from all other diseases.

Furthermore, all the immunohistochemical markers, especially CEA, are of value in exclusion of mesothelioma as sensitivity was in general high. So, positive staining for CEA and Ber-EP4 and negative staining for EMA and calretinin are reassuring that a patient does not have mesothelioma. The specificity of these markers varied and depended on the comparison group and therefore the differential diagnosis. SMRP might be of value confirming the diagnosis mesothelioma when a high cut-off-value is applied (resulting in high specificity).

Our data involved the markers used for cytological examination of pleural fluid and ascites, as well as markers used to test serum, and pleural fluid and ascites levels. To our knowledge no comprehensive systematic literature search on immunohistochemical markers in the cytological diagnosis of mesothelioma has been performed previous to this study. Recently, a meta-analysis was published on the diagnostic performance of serum SMRP only.[18] Notwithstanding large differences in the methods of data extraction, the inferences of that review were consistent with ours. Still, we come to another conclusion about the study quality. Other meta-analyses on effusion markers focused on differentiating benign from malignant diseases in general, and as such are not directly comparable with our review as our focus was to quantify the diagnostic accuracy of these markers for discriminating mesothelioma from non-mesothelioma.[19;20] Other reviews in this field did not at all perform a systematic search, and might thus be liable to selection bias in terms of included studies.[21-23]

To appreciate this systematic review, various issues should be addressed. First, the rather low quality of the eligible studies limits the conclusions about the value of markers in the diagnosis of mesothelioma. Therefore, conform to prevailing guidelines of diagnostic meta-analyses, we explicitly refrained to meta-analyse or pool the sensitivities and specificities of the individual markers. The low quality might be partly explained by including all studies with information on markers for mesothelioma regardless of their main objective. The design of most studies was exploratory, rather than confirmatory, which is illustrated by the fact that 88 markers were studies in the 82 selected papers. Exclusion of all studies with low quality scores on the QUADAS instrument would have interfered with our main objective to obtain a complete overview of markers, and was therefore not done. Furthermore, just a few studies had an acceptable quality, and only two studies had a prospective selection of consecutive patients suspected of mesothelioma. Several other studies used a prospective, consecutive patient inclusion, but selected patients on grounds of the presence of pleural effusion, rather than the initial suspicion of mesothelioma.[r2;r31;r34] Once pleural effusion is confirmed by imaging, only those patients that are still suspected of mesothelioma after

imaging are warranted for further testing for mesothelioma. The most frequently applied design was the case-control design, in a retrospective fashion. This design has been criticized for leading to biased estimates of accuracy.[9;12-16] Due to this high number of case-control studies we could not validly combine 'benign and other-malignant diseases' into one control group. Otherwise, overall sensitivity and specificity would have been strongly depended on the distribution of other-malignant and non-malignant diseases, included in these studies.
Second, reporting of study details was also poor. For example, some studies explicitly stated that they excluded paucicellular cytological samples, whereas the majority of studies provided no details about which types of other-malignant or non-malignant cases were included in the control subjects. Due to the low quality and poor reporting of study details, we could also not explore study heterogeneity.
Third, we did not assess the diagnostic value of combined markers but focused on the value of single markers instead. Pathological examination of effusion includes the use of several immunohistochemical markers. However, as studies used different combinations of markers we did not have sufficient studies to properly meta-analyse their diagnostic accuracy. Nevertheless, knowledge of the value of individual markers will certainly add to the performance of combined marker sets.
Fourth, we did not search for non published studies due to the large number of studies identified. Hence, our results may suffer from publication bias. Also, studies which did not report proportions of patients above or below a certain cut-off value could not be included in our analysis since no two-by-two table could be constructed. This mainly involved studies which showed no difference in mean or median marker levels among groups.
Finally, head to head comparisons are preferred to meta-analytically compare the diagnostic accuracy of markers. Although sufficient studies were performed to evaluate both CEA and Ber-Ep4, no robust conclusion could be drawn on their relative performance, due to the heterogeneity of the studies.

Having raised these concerns, the question remains which markers are most suitable for use in clinical practice. The aim of developing serum and cytological markers is to establish a non-invasive diagnosis of mesothelioma to prevent the already weakened patient undergoing invasive tests. In addition, the diagnosis of mesothelioma should be firm to enable a financial compensation, requiring markers to have high specificity. A major advantage of SMRP is that it can be applied by the patient's physician, whereas the use of cytological immunohistochemical markers is reserved for a pathologist. Unfortunately, the diagnostic performance of SMRP alone seems not (yet) high enough for that purpose. The specificity of cytological markers (CEA, Ber-EP4, calretinin) appears to be rather heterogeneous, potentially, due to differences in study quality, marker handling, type of antibody, type of effusion and patient and sample selection among studies. EMA will only yield a high specificity when the differential diagnosis is between mesothelioma and reactive mesothelial proliferation. However, the EMA marker was not always 100% specific across the studies.

Moreover, the value of markers, in particular immunohistochemical markers, depends on the type of mesothelioma. Sarcomatoid mesothelioma, which accounts for about 15% of all mesotheliomas, shed almost no malignant cells into the fluid making markers less useful.[24] Most immunohistochemical studies that we scrutinized included only epitheloid and biphasic mesotheliomas. Furthermore, morphology plays a major role in the decision-making process when evaluating cytological samples. Unfortunately, the majority of the studies did not consider the (added) value of immunohistochemical staining in relation to morphology.

To date, the vast majority of the studies on mesothelioma markers seem to involve rather early phase diagnostic studies (using retrospective, case-control type of designs).[25;26] It seems that the next step in studying the most promising markers, is the conduction of prospective accuracy studies in the proper target population, i.e. patients selected on their suspicion of having mesothelioma, rather than on its true presence or absence.[13;14] Subsequently, the incremental marker value of these markers beyond existing diagnostics such as patient characteristics and previous clinical tests, should be investigated.[27;28] Indeed, these prospective studies are extremely hard to perform by single institutions if the disease under study has incidences as low as that of mesothelioma. Hence, we encourage researchers and physicians to join forces to enhance the proper quantification of the diagnostic accuracy of the most promising markers for mesothelioma. Alternatively, retrospective nested case-control studies could be conducted, which are especially efficient for rare diseases and if human material is stored.[13;14;27] In these studies both cases and controls can be sampled from a single source population, typically defined by the initial presentation or suspicion of the patient. This systematic review indicated that promising markers that certainly allow for further validation are SMRP, CEA, EMA, calretinin and Ber-Ep4. Additionally, other markers might be promising which have not yet been validated in a number of studies, for example TTF-1. Finally, we encourage the improvement of reporting of diagnostic accuracy studies, following the STARD guidelines.[29;30] Only accurate quantification and reporting of the (added) value of mesothelioma markers will lead to the clinical use of the appropriate markers.

Reference List

1. Fassina A, Fedeli U, Corradin M, Da FM, Fabbris L. Accuracy and reproducibility of pleural effusion cytology. Leg Med (Tokyo) 2008; 10:20-25.
2. Fletcher SV, Clark RJ. The Portsmouth thoracoscopy experience, an evaluation of service by retrospective case note analysis. Respir Med 2007; 101:1021-1025.
3. Renshaw AA, Dean BR, Antman KH, Sugarbaker DJ, Cibas ES. The role of cytologic evaluation of pleural fluid in the diagnosis of malignant mesothelioma. Chest 1997; 111:106-109.
4. Lyons-Boudreaux V, Mody DR, Zhai J, Coffey D. Cytologic malignancy versus benignancy: how useful are the "newer" markers in body fluid cytology? Arch Pathol Lab Med 2008; 132:23-28.
5. Holloway AJ, Diyagama DS, Opeskin K, Creaney J, Robinson BW, Lake RA et al. A molecular diagnostic test for distinguishing lung adenocarcinoma from malignant mesothelioma using cells collected from pleural effusions. Clin Cancer Res 2006; 12:5129-5135.
6. Pass HI, Lott D, Lonardo F, Harbut M, Liu Z, Tang N et al. Asbestos exposure, pleural mesothelioma, and serum osteopontin levels. N Engl J Med 2005; 353:1564-1573.
7. Creaney J, Yeoman D, Demelker Y, Segal A, Musk AW, Skates SJ et al. Comparison of osteopontin, megakaryocyte potentiating factor, and mesothelin proteins as markers in the serum of patients with malignant mesothelioma. J Thorac Oncol 2008; 3:851-857.
8. Scherpereel A, Grigoriu B, Conti M, Gey T, Gregoire M, Copin MC et al. Soluble mesothelin-related peptides in the diagnosis of malignant pleural mesothelioma. Am J Respir Crit Care Med 2006; 173:1155-1160.
9. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003; 3:25.
10. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 1998; 17:873-890.
11. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58:982-990.
12. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. Ann Intern Med 2004; 140:189-202.
13. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. Clin Chem 2005; 51:1335-1341.
14. Biesheuvel CJ, Vergouwe Y, Oudega R, Hoes AW, Grobbee DE, Moons KG. Advantages of the nested case-control design in diagnostic research. BMC Med Res Methodol 2008; 8:48.
15. Mol BW, Lijmer JG, Evers JL, Bossuyt PM. Characteristics of good diagnostic studies. Semin Reprod Med 2003; 21:17-25.
16. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999; 282:1061-1066.
17. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983; 39:207-215.
18. Luo L, Shi HZ, Liang QL, Jiang J, Qin SM, Deng JM. Diagnostic value of soluble mesothelin-related peptides for malignant mesothelioma: a meta-analysis. Respir Med 2010; 104:149-156.
19. Shi HZ, Liang QL, Jiang J, Qin XJ, Yang HB. Diagnostic value of carcinoembryonic antigen in malignant pleural effusion: a meta-analysis. Respirology 2008; 13:518-527.
20. Liang QL, Shi HZ, Qin XJ, Liang XD, Jiang J, Yang HB. Diagnostic accuracy of tumour markers for malignant pleural effusion: a meta-analysis. Thorax 2008; 63:35-41.
21. Greillier L, Baas P, Welch JJ, Hasan B, Passioukov A. Biomarkers for malignant pleural mesothelioma: current status. Mol Diagn Ther 2008; 12:375-390.
22. Scherpereel A, Lee YC. Biomarkers for mesothelioma. Curr Opin Pulm Med 2007; 13:339-443.
23. Creaney J, Robinson BW. Serum and pleural fluid biomarkers for mesothelioma. Curr Opin Pulm Med 2009; 15:366-370.
24. Husain AN, Colby TV, Ordonez NG, Krausz T, Borczuk A, Cagle PT et al. Guidelines for pathologic diagnosis of malignant mesothelioma: a consensus statement from the International Mesothelioma Interest Group. Arch Pathol Lab Med 2009; 133:1317-1331.
25. Lijmer JG, Leeflang M, Bossuyt PM. Proposals for a phased evaluation of medical tests. Med Decis Making 2009; 29:E13-E21.
26. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991; 11:88-94.
27. Moons KG. Criteria for scientific evaluation of novel markers: a perspective. Clin Chem 2010; 56:537-541.
28. Riley RD, Sauerbrei W, Altman DG. Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis, and beyond. Br J Cancer 2009; 100:1219-1229.
29. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Toward complete and accurate reporting of studies of diagnostic accuracy. The STARD initiative. Am J Clin Pathol 2003; 119:18-22.
30. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem 2003; 49:7-18.

Appendix 1. The search strategy

Database	Search strategy
Pubmed	((mesothelioma[mesh] OR mesothelioma*[all]) AND ("Body Fluids"[Mesh] OR pleural effusion[Mesh] OR pericardial effusion[Mesh] OR cytology[tiab] OR cytologic*[tiab] OR serum[All] OR blood[all] OR serous[all] OR effusion[all] OR effusions[all] OR fluid*[all] OR ascites[all]) AND (biomarker[Mesh] OR biomarker*[all] OR marker*[all] OR protein[all] OR peptide[all] OR antibody[all] OR elisa[all] OR gene[all] OR genetic*[all] OR oncogene*[all] OR chromosome[all] OR chromosomal[all])) OR ((mesothelioma[mesh] OR mesothelioma*[all]) AND (immunocytochemistry[tiab] OR immunocytology[tiab]))
Embase	((mesothelioma* OR 'mesothelioma'/exp) AND [embase]/lim AND ('biomarker'/exp OR biomarker* OR marker* OR protein OR peptide OR antibody OR elisa OR gene OR genetic* OR oncogene* OR chromosome OR chromosomal) AND ('body fluids'/exp OR 'pleural effusion'/exp OR 'pericardial effusion' OR cytology:ab,ti OR cytologic*:ab,ti OR serum OR blood OR serous OR effusion OR effusions OR fluid* OR ascites)) OR ((mesothelioma* OR 'mesothelioma'/exp) AND [embase]/lim AND (immunocytochemistry:ab,ti OR immunocytology:ab,ti))

Appendix 2. The QUADAS instrument to assess methodological quality of individual diagnostic accuracy studies

Study quality was assessed using the QUADAS criteria[9], with each item scored "yes", "no", or "unclear". The items of the QUADAS tool and their interpretation are described below. In the definitions, the index test refers to the marker under study (i.e. a serum marker, effusion marker, immunohistochemical marker or genetic marker).

1a. Representative spectrum?
The participants of interest were patients with suspected mesothelioma, selected following a prospective, consecutive patient inclusion. Retrospective cohort studies and case-control studies were scored as no.

1b. Clear description of the study participants?
The description of the study participants was considered sufficient if the age distribution, female-to-male ratio and description of the disease were sufficiently described (per study group). If not, this item was scored as no. This item was supplementary to the first QUADAS criterion as a description of the characteristics is important to judge the population actually included and therefore to judge generalisability.

2. Clear description of selection criteria?
The description of the selection criteria was considered sufficient when time period and location of recruitment and setting were described, if it was clear if data collection was planned before (prospective study) or after (retrospective study) the index test and reference standard were performed, and how participants were recruited, i.e. based on presented symptoms or on the fact that the participants had received the index tests or a (specific) reference standard.

3. Is the reference standard likely to correctly classify the target condition?
This item was scored yes when diagnosis of mesothelioma was based on at least cytology or histology.

4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
This item was scored yes when the time interval between index test and reference test was less than one month in >80% of the mesothelioma patients. Unclear was utilized when this percentage could not be calculated or no information was given. When it was stated that specimens were collected at time of diagnosis without further specification this item was scored unclear as well.

5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
If all patients or a random selection of patients received verification with the reference standard then the item was scored yes, even if the reference standard was not the same (see next item) for all patients. When no information was provided about the flow of the patients, or patients were not selected consecutively, then this item was scored unclear. Case-control or retrospective cohort studies scored no on this item as this design commonly leads to partial verification bias due to a non-random selection.[12-17]

6. Did patients receive the same reference standard regardless of the index test result?
This item was scored yes when patients received the same reference standard or when the index test was performed after the reference test (as it is unlikely that the reference test will affect the performance of the index test). Unclear was used when it was uncertain if the index test was (also) performed before or after the reference test.

7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
When the results of the index test were not incorporated in the final diagnosis of all study patients the item was scored yes, or when the index test was performed after the established diagnosis. When evaluating cytologic markers and no details were provided on the cytologic markers that were used in the cytology reference standard the item was scored unclear.

8a. Was the execution of the index test described in sufficient detail to permit replication of the test?
To score yes required that the description included how antibodies or markers of genes were retrieved (or the name of the manufacturer) and type of detection system. In addition, the process of handling and preparation of samples (if cell blocks, cytospins, smears, fresh or stored samples were used) had to be included in case of immunohistochemical markers.

8b.Was the cut-off value described clearly?
This item was scored yes when a clear definition of units, cut-off point or categories of the results of the index test were provided. For immunohistochemical markers this item was scored yes when it was clear which percentage of cells, type of staining pattern and intensity were considered as positive.

9. Was the execution of the reference standard described in sufficient detail to permit its replication?
To score yes the description had to include the criteria that was used in the reference standard of mesothelioma (i.e. based on morphologic features, certain cytological or histological markers (including details on type) or electron microscope). When it was stated that diagnosis was based following published guidelines this item was scored as yes. At first, to score yes the number of mesothelioma patients that was diagnosed by either cytology or histology had to be reported, however this was almost not reported in the studies and therefore not incorporated.

10. Were the index test results interpreted without knowledge of the results of the reference standard?
To confirm that this blinding was accounted for, a clear statement in the text such as "personnel/observers who performed the biomarker assessment were blinded/unaware of the patient's diagnosis" had to be given. If there was a statement that blinding was not accounted for the item was scored as no. If no statement on blinding was given the item was scored unclear. If the index test was entirely quantitative (and required no subjective interpretation), e.g. a test using ELISA or immunoradiometric assay, then this item was scored yes.

11. Were the reference standard results interpreted without knowledge of the results of the index test?
To confirm that this blinding was accounted for, a clear statement in the text such as "personnel/observers who performed the biomarker assessment were blinded/unaware of the results of the index test" had to be given. If it was clear that blinding was not accounted for the item was scored as no. If no statement on blinding was given and the index test did not form part of the reference standard the item was scored unclear. When it was clear that the index test was performed subsequent to the diagnosis we scored this item as yes. If it was unclear whether the index was (also) performed before or after the reference standard the item was scored as unclear.

12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
When it was clear that pre-test or other clinical data were available when the index test (biomarker assay) was interpreted, then the item was scored yes. When it was stated that

observers of the biomarker assay were blinded to clinical data the item was scored with no. Unclear was used when no statement on the availability of pre-test or clinical data was provided. If the test was entirely objective (i.e. a test using ELISA or immunoradiometric assay) then this item was scored yes.

13. Were uninterpretable/ intermediate/ test results reported or other missing test results explained?
If uninterpretable, failed or intermediate results were documented or all results were available for all patients who entered the study then the item was scored yes. If it was apparent that results were missing but no explanation was given, the item was scored no. When missing results were due to the fact that specimens were not available for all participants and it was not further specified why, we scored this item as no.

Appendix 3. Reference list of studies included in the systematic review

r1 Aleman C, Manuel PJ, Segura MA, Alegre J, Esquerda A, Ruiz E et al. Pleural fluid mesothelin for the differential diagnosis of exudative pleural effusions. Med Clin (Barc) 2009 October 3;133(12):449-53.

r2 Davies HE, Sadler RS, Bielsa S, Maskell NA, Rahman NM, Davies RJ et al. The Clinical Impact and Reliability of Pleural Fluid Mesothelin in Undiagnosed Pleural Effusions. Am J Respir Crit Care Med 2009 March 19.

r3 Grigoriu B, Chahine B, Zerimech F, Gregoire M, Balduyck M, Copin MC et al. Serum mesothelin has a higher diagnostic utility than hyaluronic acid in malignant mesothelioma. Clin Biochem 2009 July;42(10-11):1046-50.

r4 Rodriguez Portal JA, Rodriguez BE, Rodriguez RD, Alfageme M, I, Quero MA, Diego RC et al. Serum levels of soluble mesothelin-related peptides in malignant and nonmalignant asbestos-related pleural disease: relation with past asbestos exposure. Cancer Epidemiol Biomarkers Prev 2009 February;18(2):646-50.

r5 Shigematsu Y, Hanagiri T, Kuroda K, Baba T, Mizukami M, Ichiki Y et al. Malignant mesothelioma-associated antigens recognized by tumor-infiltrating B cells and the clinical significance of the antibody titers. Cancer Sci 2009 April 30.

r6 Amati M, Tomasetti M, Scartozzi M, Mariotti L, Alleva R, Pignotti E et al. Profiling tumor-associated markers for early detection of malignant mesothelioma: an epidemiologic study. Cancer Epidemiol Biomarkers Prev 2008 January;17(1):163-70.

r7 Creaney J, Yeoman D, Demelker Y, Segal A, Musk AW, Skates SJ et al. Comparison of osteopontin, megakaryocyte potentiating factor, and mesothelin proteins as markers in the serum of patients with malignant mesothelioma. J Thorac Oncol 2008 August;3(8):851-7.

r8 Iwahori K, Osaki T, Serada S, Fujimoto M, Suzuki H, Kishi Y et al. Megakaryocyte potentiating factor as a tumor marker of malignant pleural mesothelioma: evaluation in comparison with mesothelin. Lung Cancer 2008 October;62(1):45-54.

r9 Pass HI, Wali A, Tang N, Ivanova A, Ivanov S, Harbut M et al. Soluble mesothelin-related peptide level elevation in mesothelioma serum and pleural effusions. Ann Thorac Surg 2008 January;85(1):265-72.

r10 Schneider J, Hoffmann H, Dienemann H, Herth FJ, Meister M, Muley T. Diagnostic and prognostic value of soluble mesothelin-related proteins in patients with malignant pleural mesothelioma in comparison with benign asbestosis and lung cancer. J Thorac Oncol 2008 November;3(11):1317-24.

r11 Creaney J, van B, I, Hof M, Segal A, Musk AW, de KN et al. Combined CA125 and mesothelin levels for the diagnosis of malignant mesothelioma. Chest 2007 October;132(4):1239-46.

r12 Creaney J, Yeoman D, Naumoff LK, Hof M, Segal A, Musk AW et al. Soluble mesothelin in effusions: a useful tool for the diagnosis of malignant mesothelioma. Thorax 2007 July;62(7):569-76.

r13 Cristaudo A, Foddis R, Vivaldi A, Guglielmi G, Dipalma N, Filiberti R et al. Clinical significance of serum mesothelin in patients with mesothelioma and lung cancer. Clin Cancer Res 2007 September 1;13(17):5076-81.

r14 Di SF, Fontana A, Loizzi M, Capotorto G, Maggiolini P, Mera E et al. Mesothelin family proteins and diagnosis of mesothelioma: analytical evaluation of an automated immunoassay and preliminary clinical results. Clin Chem Lab Med 2007;45(5):634-8.

r15 Grigoriu BD, Scherpereel A, Devos P, Chahine B, Letourneux M, Lebailly P et al. Utility of osteopontin and serum mesothelin in malignant pleural mesothelioma diagnosis and prognosis assessment. Clin Cancer Res 2007 May 15;13(10):2928-35.

r16 Shiomi K, Hagiwara Y, Sonoue K, Segawa T, Miyashita K, Maeda M et al. Sensitive and specific new enzyme-linked immunosorbent assay for N-ERC/mesothelin increases its potential as a useful serum tumor marker for mesothelioma. Clin Cancer Res 2008 March 1;14(5):1431-7.

r17 van den Heuvel MM, Korse CM, Bonfrer JM, Baas P. Non-invasive diagnosis of pleural malignancies: the role of tumour markers. Lung Cancer 2008 March;59(3):350-4.

r18 Welker L, Muller M, Holz O, Vollmer E, Magnussen H, Jorres RA. Cytological diagnosis of malignant mesothelioma - Improvement by additional analysis of hyaluronic acid in pleural effusions. Virchows Arch 2007;450(4):455-61.

r19 Onda M, Nagata S, Ho M, Bera TK, Hassan R, Alexander RH et al. Megakaryocyte potentiation factor cleaved from mesothelin precursor is a useful tumor marker in the serum of patients with mesothelioma. Clin Cancer Res 2006 July 15;12(14 Pt 1):4225-31.

r20 Filiberti R, Marroni P, Neri M, Ardizzoni A, Betta PG, Cafferata MA et al. Serum PDGF-AB in pleural mesothelioma. Tumour Biol 2005 September;26(5):221-6.

r21 Pass HI, Lott D, Lonardo F, Harbut M, Liu Z, Tang N et al. Asbestos exposure, pleural mesothelioma, and serum osteopontin levels. N Engl J Med 2005 October 13;353(15):1564-73.

r22 Scherpereel A, Grigoriu B, Conti M, Gey T, Gregoire M, Copin MC et al. Soluble mesothelin-related peptides in the diagnosis of malignant pleural mesothelioma. Am J Respir Crit Care Med 2006 May 15;173(10):1155-60.

r23 Neri M, Betta P, Marroni P, Filiberti R, Cafferata M, Mereu C et al. Serum anti-p53 autoantibodies in pleural malignant mesothelioma, lung cancer and non-neoplastic lung diseases. Lung Cancer 2003 February;39(2):165-72.

r24 Villena V, Lopez-Encuentra A, Echave-Sustaeta J, Martin-Escribano P, Ortuno-de-Solo B, Estenoz-Alfaro J. Diagnostic value of CA 549 in pleural fluid. Comparison with CEA, CA 15.3 and CA 72.4. Lung Cancer 2003 June;40(3):289-94.

r25 Creaney J, McLaren BM, Stevenson S, Musk AW, de KN, Robinson BW et al. p53 autoantibodies in patients with malignant mesothelioma: stability through disease progression. Br J Cancer 2001 January 5;84(1):52-6.

r26 Paganuzzi M, Onetto M, Marroni P, Filiberti R, Tassara E, Parodi S et al. Diagnostic value of CYFRA 21-1 tumor marker and CEA in pleural effusion due to mesothelioma. Chest 2001 April;119(4):1138-42.

r27 Fuhrman C, Duche JC, Chouaid C, Abd A, I, Atassi K, Monnet I et al. Use of tumor markers for differential diagnosis of mesothelioma and secondary pleural malignancies. Clin Biochem 2000 July;33(5):405-10.

r28 Alatas F, Alatas O, Metintas M, Colak O, Harmanci E, Demir S. Diagnostic value of CEA, CA 15-3, CA 19-9, CYFRA 21-1, NSE and TSA assay in pleural effusions. Lung Cancer 2001 January;31(1):9-16.

r29 Miedouge M, Rouzaud P, Salama G, Pujazon MC, Vincent C, Mauduyt MA et al. Evaluation of seven tumour markers in pleural fluid for the diagnosis of malignant effusions. Br J Cancer 1999 November;81(6):1059-65.

r30 Nisman B, Barak V, Heching N, Kramer M, Reinus C, Lafair J. Cytokeratin markers in malignant pleural mesothelioma. Cancer Detect Prev 1998;22(5):416-21.

r31 Atagi S, Ogawara M, Kawahara M, Sakatani M, Furuse K, Ueda E et al. Utility of hyaluronic acid in pleural fluid for differential diagnosis of pleural effusions: likelihood ratios for malignant mesothelioma. Jpn J Clin Oncol 1997 October;27(5):293-7.

r32 Ebert W, Hoppe M, Muley T, Drings P. Monitoring of therapy in inoperable lung cancer patients by measurement of CYFRA 21-1, TPA- TP CEA, and NSE. Anticancer Res 1997 July;17(4B):2875-8.

r33 Shijubo N, Honda Y, Fujishima T, Takahashi H, Kodama T, Kuroki Y et al. Lung surfactant protein-A and carcinoembryonic antigen in pleural effusions due to lung adenocarcinoma and malignant mesothelioma. Eur Respir J 1995 March;8(3):403-6.

r34 Villena V, Lopez-Encuentra A, Echave-Sustaeta J, Martin-Escribano P, Ortuno-de-Solo B, Estenoz-Alfaro J. Diagnostic value of CA 72-4, carcinoembryonic antigen, CA 15-3, and CA 19-9 assay in pleural fluid. A study of 207 patients. Cancer 1996 August 15;78(4):736-40.

r35 Whitaker D, Shilkin KB, Stuckey M, Nieuwhof WN. Pleural fluid CEA levels in the diagnosis of malignant mesothelioma. Pathology 1986 July;18(3):328-9.

r36 Faravelli B, D'Amore E, Nosenzo M, Betta PG, Donna A. Carcinoembryonic antigen in pleural effusions. Diagnostic value in malignant mesothelioma. Cancer 1984 March 1;53(5):1194-7.

r37 Shen J, Pinkus GS, Deshpande V, Cibas ES. Usefulness of EMA, GLUT-1, and XIAP for the cytologic diagnosis of malignant mesothelioma in body cavity fluids. Am J Clin Pathol 2009 April;131(4):516-23.

r38 Slipicevic A, Oy GF, Askildt IC, Holth A, Hellesylt E, Florenes VA et al. Diagnostic and prognostic role of the insulin growth factor pathway members insulin-like growth factor-II and insulin-like growth factor binding protein-3 in serous effusions. Hum Pathol 2009 April;40(4):527-37.

r39 Yuan Y, Nymoen DA, Stavnes HT, Rosnes AK, Bjorang O, Wu C et al. Tenascin-X is a novel diagnostic marker of malignant mesothelioma. Am J Surg Pathol 2009 November;33(11):1673-82.

r40 Bhalla R, Siddiqui MT, Mandich D, Cartun RW, Fiel-Gan MD, Nassar A et al. Diagnostic utility of D2-40 and podoplanin in effusion cell blocks. Diagn Cytopathol 2007 June;35(6):342-7.

r41 Facchetti F, Lonardi S, Gentili F, Bercich L, Falchetti M, Tardanico R et al. Claudin 4 identifies a wide spectrum of epithelial neoplasms and represents a very useful marker for carcinoma versus mesothelioma diagnosis in pleural and peritoneal biopsies and effusions. Virchows Arch 2007 September;451(3):669-80.

r42 Grefte JM, de Wilde PC, Salet-van de Pol MR, Tomassen M, Raaymakers-van Geloof WL, Bulten J. Improved identification of malignant cells in serous effusions using a small, robust panel of antibodies on paraffin-embedded cell suspensions. Acta Cytol 2008 January;52(1):35-44.

r43 Kleinberg L, Holth A, Fridman E, Schwartz I, Shih I, Davidson B. The diagnostic role of claudins in serous effusions. Am J Clin Pathol 2007 June;127(6):928-37.

r44 Pu RT, Pang Y, Michael CW. Utility of WT-1, p63, MOC31, mesothelin, and cytokeratin (K903 and CK5/6) immunostains in differentiating adenocarcinoma, squamous cell carcinoma, and malignant mesothelioma in effusions. Diagn Cytopathol 2008 January;36(1):20-5.

r45 Shield PW, Koivurinne K. The value of calretinin and cytokeratin 5/6 as markers for mesothelioma in cell block preparations of serous effusions. Cytopathology 2008 August;19(4):218-23.

r46 Aerts JG, Delahaye M, van der Kwast TH, Davidson B, Hoogsteden HC, van Meerbeeck JP. The high post-test probability of a cytological examination renders further investigations to establish a diagnosis of epithelial malignant pleural mesothelioma redundant. Diagn Cytopathol 2006 August;34(8):523-7.

r47 Bassarova AV, Nesland JM, Davidson B. D2-40 is not a specific marker for cells of mesothelial origin in serous effusions. Am J Surg Pathol 2006 July;30(7):878-82.

r48 Li Q, Bavikatty N, Michael CW. The role of immunohistochemistry in distinguishing squamous cell carcinoma from mesothelioma and adenocarcinoma in pleural effusion. Semin Diagn Pathol 2006 February;23(1):15-9.

r49 Saad RS, Lindner JL, Lin X, Liu YL, Silverman JF. The diagnostic utility of D2-40 for malignant mesothelioma versus pulmonary carcinoma with pleural involvement. Diagn Cytopathol 2006 December;34(12):801-6.

r50 Sivertsen S, Berner A, Michael CW, Bedrossian C, Davidson B. Cadherin expression in ovarian carcinoma and malignant mesothelioma cell effusions. Acta Cytol 2006 November;50(6):603-7.

r51 Afify AM, Stern R, Michael CW. Differentiation of mesothelioma from adenocarcinoma in serous effusions: the role of hyaluronic acid and CD44 localization. Diagn Cytopathol 2005 March;32(3):145-50.

r52 Hecht JL, Pinkus JL, Pinkus GS. Monoclonal antibody MOC-31 reactivity as a marker for adenocarcinoma in cytologic preparations. Cancer 2006 February 25;108(1):56-9.

r53 Saad RS, Cho P, Liu YL, Silverman JF. The value of epithelial membrane antigen expression in separating benign mesothelial proliferation from malignant mesothelioma: a comparative study. Diagn Cytopathol 2005 March;32(3):156-9.

r54 Saqi A, Yun SS, Yu GH, Alexis D, Taub RN, Powell CA et al. Utility of CD138 (syndecan-1) in distinguishing carcinomas from mesotheliomas. Diagn Cytopathol 2005 August;33(2):65-70.

r55 Schonherr A, Bayer M, Bocking A. Diagnostic and prognostic value of Ki67 proliferation fraction in serous effusions. Cell Oncol 2004;26(1-2):57-62.

r56 Afify AM, al-Khafaji BM. Diagnostic utility of thyroid transcription factor-1 expression in adenocarcinomas presenting in serous fluids. Acta Cytol 2002 July;46(4):675-8.

r57 Afify AM, al-Khafaji BM, Paulino AF, Davila RM. Diagnostic use of muscle markers in the cytologic evaluation of serous fluids. Appl Immunohistochem Mol Morphol 2002 June;10(2):178-82.

r58 Davidson B, Nielsen S, Christensen J, Asschenfeldt P, Berner A, Risberg B et al. The role of desmin and N-cadherin in effusion cytology: a comparative study using established markers of mesothelial and epithelial cells. Am J Surg Pathol 2001 November;25(11):1405-12.

r59 Hecht JL, Lee BH, Pinkus JL, Pinkus GS. The value of Wilms tumor susceptibility gene 1 in cytologic preparations as a marker for malignant mesothelioma. Cancer 2002 April 25;96(2):105-9.

r60 Hecht JL, Pinkus JL, Weinstein LJ, Pinkus GS. The value of thyroid transcription factor-1 in cytologic preparations as a marker for metastatic adenocarcinoma of lung origin. Am J Clin Pathol 2001 October;116(4):483-8.

r61 Simsir A, Fetsch P, Bedrossian CW, Ioffe OB, Abati A. Absence of SV-40 large T antigen (Tag) in malignant mesothelioma effusions: an immunocytochemical study. Diagn Cytopathol 2001 October;25(4):203-7.

r62 Wieczorek TJ, Krane JF. Diagnostic utility of calretinin immunohistochemistry in cytologic cell block preparations. Cancer 2000 October 25;90(5):312-9.

r63 Dejmek A, Hjerpe A. Reactivity of six antibodies in effusions of mesothelioma, adenocarcinoma and mesotheliosis: stepwise logistic regression analysis. Cytopathology 2000 February;11(1):8-17.

r64 Motherby H, Kube M, Friedrichs N, Nadjari B, Knops K, Donner A et al. Immunocytochemistry and DNA-image cytometry in diagnostic effusion cytology I. Prevalence of markers in tumour cell positive and negative smears. Anal Cell Pathol 1999;19(1):7-20.

r65 Simsir A, Fetsch P, Mehta D, Zakowski M, Abati A. E-cadherin, N-cadherin, and calretinin in pleural effusions: the good, the bad, the worthless. Diagn Cytopathol 1999 March;20(3):125-30.

r66 Ascoli V, Carnovale-Scalzo C, Taccogna S, Nardi F. Utility of HBME-1 immunostaining in serous effusions. Cytopathology 1997 October;8(5):328-35.

r67 Delahaye M, van der HF, van der Kwast TH. Complementary value of five carcinoma markers for the diagnosis of malignant mesothelioma, adenocarcinoma metastasis, and reactive mesothelium in serous effusions. Diagn Cytopathol 1997 August;17(2):115-20.

r68 Ascoli V, Scalzo CC, Taccogna S, Nardi F. The diagnostic value of thrombomodulin immunolocalization in serous effusions. Arch Pathol Lab Med 1995 December;119(12):1136-40.

r69 Baars JH, De Ruijter JLM, Smedts F, Van Niekerk CC, Poels LG, Seldenrijk CA et al. The applicability of a keratin 7 monoclonal antibody in routinely papanicolaou-stained cytologic specimens for the differential diagnosis of carcinomas. Am J Clin Pathol 1994;101(3):257-61.

r70 Donna A, Betta PG, Bellingeri D, Tallarida F, Pavesi M, Pastormerlo M. Cytologic diagnosis of malignant mesothelioma in serous effusions using an antimesothelial-cell antibody. Diagn Cytopathol 1992;8(4):361-5.

r71 Betta PG, Pavesi M, Pastormerlo M, Tallarida F, Bellingeri D, Bocca R. Use of monoclonal antibody B72.3 as a marker of metastatic carcinoma cells in neoplastic effusions. Pathologica 1991 January;83(1083):99-104.

r72 Delahaye M, Hoogsteden HC, van der Kwast TH. Immunocytochemistry of malignant mesothelioma: OV632 as a marker of malignant mesothelioma. J Pathol 1991 October;165(2):137-43.

r73 Kuhlmann L, Berghauser KH, Schaffer R. Distinction of mesothelioma from carcinoma in pleural effusions. An immunocytochemical study on routinely processed cytoblock preparations. Pathol Res Pract 1991 May;187(4):467-71.

r74 Linari A, Bussolati G. Evaluation of impact of immunocytochemical techniques in cytological diagnosis of neoplastic effusions. J Clin Pathol 1989 November;42(11):1184-9.

r75 Cibas ES, Corson JM, Pinkus GS. The distinction of adenocarcinoma from malignant mesothelioma in cell blocks of effusions: the role of routine mucin histochemistry and immunohistochemical assessment of carcinoembryonic antigen, keratin proteins, epithelial membrane antigen, and milk fat globule-derived antigen. Hum Pathol 1987 January;18(1):67-74.

r76 Ghosh AK, Butler EB. Immunocytological staining reactions of anti-carcinoembryonic antigen, Ca, and anti-human milk fat globule monoclonal antibodies on benign and malignant exfoliated mesothelial cells. J Clin Pathol 1987 December;40(12):1424-7.

r77 Walts AE, Said JW, Shintaku IP, Sassoon AF, Banks-Schlegel S. Keratins of different molecular weight in exfoliated mesothelial and adenocarcinoma cells--an aid to cell identification. Am J Clin Pathol 1984 April;81(4):442-6.

r78 Illei PB, Ladanyi M, Rusch VW, Zakowski MF. The use of CDKN2A deletion as a diagnostic marker for malignant mesothelioma in body cavity effusions. Cancer 2003 February 25;99(1):51-6.

r79 Flores-Staino C, rai-Ramqvist E, Dobra K, Hjerpe A. Adaptation of a commercial fluorescent in situ hybridization test to the diagnosis of malignant cells in effusions. Lung Cancer 2009 June 10.

r80 Onofre FB, Onofre AS, Pomjanski N, Buckstegge B, Grote HJ, Bocking A. 9p21 Deletion in the diagnosis of malignant mesothelioma in serous effusions additional to immunocytochemistry, DNA-ICM, and AgNOR analysis. Cancer 2008 June 25;114(3):204-15.

r81 Creaney J, Segal A, Sterrett G, Platten MA, Baker E, Murch AR et al. Overexpression and altered glycosylation of MUC1 in malignant mesothelioma. Br J Cancer 2008 May 6;98(9):1562-9.

r82 Dejmek A, Hjerpe A. The combination of CEA, EMA, and BerEp4 and hyaluronan analysis specifically identifies 79% of all histologically verified mesotheliomas causing an effusion. Diagn Cytopathol 2005 March;32(3):160-6.

Appendix 4. Sensitivity and specificity of markers

Table 2.1: Reported sensitivity and specificity of serum markers per study, stratified by type of marker

No.	Marker	1st Author-year	Type of mesthelioma	in comparison with					
				Malignancy		Non-malignancy		Malignancy and non-malignancy combined	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
1.1	SMRP	Rodriquez Portal - 2009[r4]	pleural			72 (26/10)	72		
1.2	SMRP	Schneider - 2008[r10]	pleural	30 (30/70)	95 (132/7)	42 (42/58)	95 (71/4)		
1.3	SMRP	Creaney - 2008[r7]	pleural	73 (48/18)	47 (14/16)	73 (48/18)	89 (62/8)		
1.4	SMRP	Iwahori - 2008[r8]	pleural	59 (16/11)	79 (65/17)	59 (16/11)	94 (44/3)		
1.5	SMRP	Amati - 2008[r6 ^]	pleural			73 (16/6)	90 (85/9)		
1.6	SMRP	Pass - 2008[r9]	pleural	79 (71/19)	76	60 (54/36)	89 (59/7)		
1.7	SMRP	Van den Heuvel - 2007[r17]	pleural	60 (44/29)	79 (84/22)				
1.8	SMRP	Cristaudo - 2007[r13]	pleural	27 (29/78)	95	46 (49/58)	94		
1.9	SMRP	Di Serio - 2007[r14]	pleural			67 (16/8)	92 (85/7)		
1.10	SMRP	Scherpereel - 2005[r22]	pleural	58 (35/25)	73 (22/8)	80 (48/12)	83 (19/4)		
2.1	CEA*	Van den Heuvel - 2007[r17]	pleural	90 (66/7)	52 (55/51)				
2.2	CEA*	Fuhrman - 2000[r27 ^]	pleural	50 (13/13)	88 (23/3)				
2.3	CEA*	Alatas - 1999[r28]	pleural	55 (11/9)	83 (20/4)				
2.4	CEA*	Nisman - 1998[r30]	pleural	100 (14/0)	48 (39/42)	100 (14/0)	0 (0/90)		
2.5	CEA*	Ebert - 1997[r32 ^]	pleural	88 (29/4)	39 (59/91)	88 (29/4)	2 (4/182)		
3.1	CYFRA21-1	Van den Heuvel - 2007[r17]	pleural	66 (48/25)	33 (35/71)				
3.2	CYFRA21-1	Alatas - 1999[r28]	pleural			50 (10/10)	74 (29/10)		
3.3	CYFRA21-1	Nisman - 1998[r30]	pleural	50 (7/7)	56 (45/36)	50 (7/7)	93 (90/7)		
3.4	CYFRA21-1	Ebert - 1997[r32 ^]	pleural	36 (12/21)	57 (85/65)	36 (12/21)	94		
4.1	MPF	Creaney - 2008[r7]	pleural	32 (21/45)	93 (28/2)	32 (21/45)	96 (67/3)		
4.2	MPF	Iwahori - 2008[r8]	pleural	74 (20/7)	84 (69/13)	74 (20/7)	98 (46/1)		
4.3	MPF	Shiomi - 2007[r16 ^]	pleural					72 (28/11)	93
4.4	MPF	Onda - 2006[r19]	pleural, peritoneal			91 (51/5)	100 (70/0)		
5.1	osteopontin	Creaney - 2008[r7]	pleural	45 (30/36)	43 (13/17)	45 (30/36)	87 (61/9)		
5.2	osteopontin	Grigoriu - 2007[r15 ^]	pleural			60 (56/38)	80 (90/22)		

Table 2.1: Reported sensitivity and specificity of serum markers per study, stratified by type of marker *(Continued)*

No.	Marker	1st Author-year	Type of mesthelioma	in comparison with					
				Malignancy		Non-malignancy		Malignancy and non-malignancy combined	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
5.3	osteopontin	Pass - 2005[r21]	pleural			78 (59/17)	86 (59/10)		
6.1	CA15-3	Creaney - 2008[r81]	pleural			35 (17/32)	85 (56/10)		
6.2	CA15-3	Alatas - 1999[r28]	pleural			80 (16/4)	56 (20/16)		
7.1	NSE*	Alatas - 1999[r28]	pleural	70 (14/6)	63 (15/9)				
7.2	NSE*	Ebert - 1997[r32 ^]	pleural	88 (29/4)	17	88 (29/4)	2 (4/182)		
8.1	p53	Neri - 2003[r23]	pleural	7 (2/28)	83 (40/8)	7 (2/28)	98 (104/2)		
8.2	p53	Creaney - 2001[r25 ^]	unknown			7 (6/82)	94 (97/6)		
9.1	TPS	Nisman - 1998[r30]	pleural	64 (9/5)	65 (53/28)	64 (9/5)	91 (90/9)		
9.2	TPS	Ebert - 1997[r32 ^]	pleural	36 (12/21)	85	36 (12/21)	94		
10	80HdG	Amati - 2008[r6 ^]	pleural			18 (4/18)	90 (85/9)		
11	bFGF	Amati - 2008[r6 ^]	pleural			45 (10/12)	90 (85/9)		
12	CA125	Creaney - 2007[r11]	pleural			42 (49/68)	78 (91/25)		
13	Gene-X	Shigematsu - 2009[r5]	pleural	56 (10/8)	100 (63/0)	56 (10/8)	100 (25/0)		
14	HA	Grigoriu - 2009[r3]	pleural	26 (20/56)	76 (25/8)	26 (20/56)	96 (26/1)		
15	HGF	Amati - 2008[r6 ^]	pleural			36 (8/14)	90 (85/9)		
16	PDGF-AB	Filiberti - 2005[r20]	pleural	43 (40/53)	70 (23/10)	43 (40/53)	82 (42/9)		
17	PDGFß	Amati - 2008[r6 ^]	pleural			45 (10/12)	90 (85/9)		
18	THBS-2	Shigematsu - 2009[r5]	pleural	89 (16/2)	100 (63/0)	89 (16/2)	92 (23/2)		
19	TPA-M*	Ebert - 1997[r32 ^]	pleural	76 (25/8)	39 (59/91)	76 (25/8)	4 (8/182)		
20	TSA	Alatas - 1999[r28]	pleural	40 (8/12)	25 (6/18)				
21	VEGFß	Amati - 2008[r6 ^]	pleural			59 (13/9)	90 (85/9)		

*sensitivity and specificity were calculated as follows: the number of mesothelioma patients below the cut-off value was defined as TP and those above the cut-off value as FN whereas the number of non-mesothelioma patients below the cut-off value was defined as FN and those above the cut-off value as TN; ^(r6) values of the asbestos exposed group were used to construct a two-by-two table because this was the largest comparison group; ^(r15) we estimated values from a ROC curve using a noticeable cut-off value that corresponded to a specificity of 80%; ^(r16) values from Figure 4A were extracted to construct the two-by-two table because it included the largest comparison group; ^(r25) values were additionally extracted from figure 2 to construct a two-by-two table for the comparison of mesothelima to non-malignancy; ^(r27) for serum CEA, we estimated values from a ROC curve, using a noticeable cut-off value of 1 ng/ml; ^(r32) values were estimated from figure 1 to construct a two-by-two table.

Table 2.2: Reported sensitivity and specificity of effusion markers per study, stratified by type marker

No.	Marker	1st Author-year	Type of effusion(s)	in comparison with					
				Malignancy		Non-malignancy		Malignancy and non-malignancy combined	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
1.1	CEA*	Villena - 2003[r24]	pleural	100 (20/0)	43 (35/46)	100 (20/0)	0 (0/151)		
1.2	CEA*	Paganuzzi - 2001[r26]	pleural	97 (31/1)	53 (21/19)	97 (31/1)	9 (3/31)		
1.3	CEA*	Fuhrman - 2000[r27]	pleural	100 (29/0)	76 (25/8)				
1.4	CEA*	Alatas - 1999[r28]	pleural	90 (18/2)	88 (21/3)				
1.5	CEA*	Miedouge - 1999[r29]	pleural	100 (11/0)	63 (129/75)	100 (11/0)	1 (1/120)		
1.6	CEA*	Atagi - 1997[r31]	pleural	77 (10/3)	68 (17/8)				
1.7	CEA*	Villena - 1995[r34]	pleural	100 (10/0)	51 (28/27)	100 (10/0)	0 (0/142)		
1.8	CEA*	Shijubo - 1995[r33]	pleural	100 (10/0)	68 (53/25)				
1.9	CEA*	Whitaker - 1986[r35]	pleural	100 (20/0)	70 (14/6)	100 (20/0)	0 (0/20)		
1.10	CEA*	Fravelli - 1984[r36]	pleural	100 (26/0)	68 (77/37)	100 (26/0)	33 (24/49)		
2.1	CA15-3	Creaney - 2008[r81]	pleural	38 (20/32)	76 (19/6)	38 (20/32)	100 (30/0)		
2.2	CA15-3	Villena - 2003[r24]	pleural	30 (6/14)	56 (45/36)	30 (6/14)	100 (151/0)		
2.3	CA15-3	Miedouge - 1999[r29]	pleural	45 (5/6)	35 (72/132)	45 (5/6)	99 (120/1)		
2.4	CA15-3	Alatas - 1999[r28]	pleural			90 (18/2)	93 (28/2)		
2.5	CA15-3	Villena - 1995[r34]	pleural	70 (7/3)	47 (26/29)	70 (7/3)	100 (142/0)		
3.1	HA	Grigoriu - 2009[r3]	pleural					64 (49/27)	97 (58/2)
3.2	HA	Welker - 2007[r18 ^]	pleural	88 (63/9)	99 (99/1)	88 (63/9)	97 (87/3)		
3.3	HA	Dejmek - 2005[r82]	unknown	35 (20/37)	100 (73/0)	35 (20/37)	100 (36/0)		
3.4	HA	Fuhrman - 2000[r27]	pleural	32 (12/25)	95 (36/2)				
3.5	HA	Atagi - 1997[r31]	pleural					37 (7/12)	99 (79/1)
4.1	SMRP	Davies - 2009[r2]	pleural	71 (17/7)	81 (54/13)	71 (17/7)	97 (73/2)		
4.2	SMRP	Aleman - 2009[r1]	pleural					56 (10/8)	92 (46/4)
4.3	SMRP	Pass - 2008[r9]	pleural	76 (34/11)	44 (16/20)	76 (34/11)	83 (25/5)		
4.4	SMRP	Creaney - 2007[r12 ^]	pleural, peritoneal	68 (40/19)	83 (58/12)	68 (40/19)	98 (88/2)		

Table 2.2: Reported sensitivity and specificity of effusion markers per study, stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	in comparison with					
				Malignancy		Non-malignancy		Malignancy and non-malignancy combined	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
4.5	SMRP	Scherpereel - 2005[r22]	pleural	58 (25/18)	93 (26/2)	77 (33/10)	24 (5/16)		
5.1	CA19-9*	Alatas - 1999[r28]	pleural	70 (14/6)	54 (13/11)				
5.2	CA19-9*	Miedouge - 1999[r29]	pleural	100 (11/0)	22 (45/159)	100 (11/0)	1 (1/120)		
5.3	CA19-9*	Villena - 1995[r34]	pleural	100 (10/0)	24 (13/42)	100 (10/0)	0 (0/142)		
6.1	CA72-4*	Villena - 2003[r24]	pleural	100 (20/0)	37 (30/51)	100 (20/0)	0 (0/151)		
6.2	CA72-4*	Miedouge - 1999[r29]	pleural	100 (11/0)	72 (147/57)	100 (11/0)	1 (1/120)		
6.3	CA72-4*	Villena - 1995[r34]	pleural	90 (9/1)	58 (32/23)	90 (9/1)	2 (3/139)		
7.1	CYFRA21-1	Paganuzzi - 2001[r26]	pleural	88 (28/4)	33 (13/27)	88 (28/4)	79 (27/7)		
7.2	CYFRA21-1	Alatas - 1999[r28]	pleural			90 (18/2)	90 (27/3)		
7.3	CYFRA21-1	Miedouge - 1999[r29]	pleural	55 (6/5)	55 (112/92)	55 (6/5)	99 (120/1)		
8.1	NSE*	Alatas - 1999[r28]	pleural	80 (16/4)	63 (15/9)				
8.2	NSE*	Miedouge - 1999[r29]	pleural	91 (10/1)	19 (38/166)	91 (10/1)	2 (3/118)		
9	CA549*	Villena - 2003[r24]	pleural	65 (13/7)	46 (37/44)	65 (13/7)	0 (0/151)		
10	SCC*	Miedouge - 1999[r29]	pleural	100 (11/0)	6 (12/192)	100 (11/0)	1 (1/120)		
11	SP-A*	Shijubo - 1995[r33]	pleural	100 (10/0)	47 (37/41)				
12	TSA	Alatas - 1999[r28]	pleural	90 (18/2)	50 (12/12)				

*sensitivity and specificity were calculated as follows: the number of mesothelioma patients below the cut-off value was defined as TP and those above the cut-off value as FN whereas the number of non-mesothelioma patients below the cut-off value was defined as FN and those above the cut-off value as TN; ^(r12) values of pleural and peritoneal effusions were combined; ^(r18) combining the results of table 1 and figure 2 and using a cut-off value of 100 mg/l, we could construct a separate two-by-two table for the comparison of mesothelioma to non-malignancy and malignancy.

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
1.1	EMA	Shen - 2009[r37]	pleural, peritoneal			86 (30/5)	87 (33/5)
1.2	EMA (any staining)	Yuan - 2009[r39 ^]	pleural, peritoneal, pericardial	97 (35/1)	0 (0/94)		
1.2	EMA (membranous staining)	Yuan - 2009[r39]	pleural, peritoneal, pericardial	92 (33/3)	100 (94/0)		
1.3	EMA (E29)	Creaney - 2008[r81 ^]	pleural			84 (16/3)	93 (14/1)
1.3	EMA (Mc5)	Creaney - 2008[r81 ^]	pleural			100 (20/0)	0 (0/14)
1.4	EMA (any staining)	Grefte - 2007[r42]	pleural, peritoneal	100 (11/0)	0 (0/12)	100 (11/0)	91 (10/1)
1.4	EMA (membranous staining)	Grefte - 2007[r42 ^]	pleural, peritoneal	91 (10/1)	100 (12/0)	91 (10/1)	91 (10/1)
1.5	EMA	Aerts - 2006[r46]	pleural	86 (12/2)	0 (0/12)	86 (12/2)	100 (13/0)
1.6	EMA (E29)	Saad - 2005[r53]	pleural			75 (15/5)	100 (20/0)
1.6	EMA (Mc5)	Saad - 2005[r53]	pleural			70 (14/6)	60 (12/8)
1.7	EMA (any staining)	Dejmek - 2005[r82]	unknown	73 (40/15)	8 (11/121)		
1.7	EMA (membranous staining)	Dejmek - 2005[r82]	unknown	58 (32/23)	99 (123/1)		
1.8	EMA	Motherby - 1999[r64]	pleural, peritoneal, pericardial, cul de sac	100 (14/0)	2 [#] (2/85)	100 (14/0)	71 [$] (37/15)
1.9	EMA	Ascoli - 1995[r68 ^]	pleural, peritoneal, hydrocoele, synovial	100 (33/0)	9 (13/139)	100 (33/0)	91 [$] (32/3)
1.10	EMA	Delahaye - 1991[r72]	pleural, peritoneal	83 (20/4)	0 (0/31)	83 (20/4)	100 (20/0)
1.11	EMA (any staining)	Cibas - 1987[r75]	pleural, peritoneal	95 (19/1)	5 (2/37)		
1.11	EMA (membranous staining)	Cibas - 1987[r75 ^]	pleural, peritoneal	55 (11/9)	82 (32/7)		
2.1	Ber-EP4*	Yuan - 2009[r39]	pleural, peritoneal, pericardial	64 (23/13)	98 (92/2)		
2.2	Ber-EP4*	Botelho - 2008[r80]	pleural, peritoneal	84 (27/5)	87 (27/4)	84 (27/5)	0 (0/39)
2.3	Ber-EP4*	Grefte - 2007[r42]	pleural, peritoneal	100 (11/0)	92 (11/1)	100 (11/0)	0 (0/11)
2.4	Ber-EP4*	Aerts - 2006[r46]	pleural	93 (13/1)	100 (12/0)	93 (13/1)	0 (0/13)
2.5	Ber-EP4*	Li - 2006[r48]	pleural	100 (12/0)	48 (10/11)		
2.6	Ber-EP4*	Dejmek - 2005[r82]	unknown	84 (46/9)	56 (77/61)		

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
2.7	Ber-EP4*	Motherby - 1999[r64]	pleural, peritoneal, pericardial, cul de sac	100 (14/0)	95 # (83/4)	100 (14/0)	0 $ (0/53)
2.8	Ber-EP4*	Delahaye - 1997[r67]	unknown	98 (40/1)	78 (69/19)	98 (40/1)	0 (0/25)
2.9	Ber-EP4*	Ascoli - 1995[r68] ^	pleural, peritoneal, hydrocoele, synovial	100 (33/0)	95 (145/7)	100 (33/0)	0 $ (0/35)
2.10	Ber-EP4*	Kuhlman - 1991[r73]	pleural	85 (17/3)	75 (15/5)	100 (20/0)	0 $ (0/20)
3.1	CEA*	Grefte - 2007[r42]	pleural, peritoneal	100 (11/0)	58 (7/5)	100 (11/0)	0 (0/11)
3.2	CEA*	Aerts - 2006[r46]	pleural	93 (13/1)	92 (11/1)	93 (13/1)	0 (0/13)
3.3	CEA*	Li - 2006[r48]	pleural	100 (12/0)	57 (12/9)		
3.4	CEA*	Dejmek - 2005[r82]	unknown	98 (49/1)	68 (93/44)		
3.5	CEA*	Davidson - 2001[r58]	pleural, peritoneal	100 (12/0)	41 (40/58)	100 (12/0)	0 (0/56)
3.6	CEA*	Motherby - 1999[r64]	pleural, peritoneal, pericardial, cul de sac	100 (14/0)	49 # (42/43)	100 (14/0)	8 $ (4/48)
3.7	CEA*	Delahaye - 1997[r67]	unknown	100 (41/0)	55 (48/40)	100 (41/0)	0 (0/25)
3.8	CEA*	Kuhlman - 1991[r73]	pleural	90 (18/2)	50 (10/10)	100 (20/0)	0 $ (0/20)
3.9	CEA*	Ghosh - 1987[r76]	pleural, peritoneal	92 (11/1)	75 (6/2)	92 (11/1)	0 (0/5)
3.10	CEA*	Cibas - 1987[r75]	pleural, peritoneal	100 (20/0)	72 (28/11)		
4.1	calretinin	Yuan - 2009[r39]	pleural, peritoneal, pericardial	97 (35/1)	79 (74/20)		
4.2	calretinin	Botelho - 2008[r80]	pleural, peritoneal	100 (32/0)	100 (31/0)	100 (32/0)	5 (2/37)
4.3	calretinin	Grefte - 2007[r42]	pleural, peritoneal	91 (10/1)	100 (12/0)	91 (10/1)	0 (0/11)
4.4	calretinin	Shield - 2007[r45]	pleural, peritoneal	97 (33/1)	97 (65/2)		
4.5	calretinin	Bhalla - 2007[r40]	pleural, peritoneal	100 (10/0)	100 (10/0)	100 (10/0)	0 (0/20)
4.6	calretinin	Saad - 2006[r49]	pleural	85 (17/3)	55 (11/9)		
4.7	calretinin	Li - 2006[r48]	pleural	100 (12/0)	100 (21/0)		
4.8	calretinin	Wieczorek - 2000[r62] ^	pleural, peritoneal	100 (29/0)	62 (24/15)		
4.9	calretinin	Simir - 1999[r65]	pleural	58 (15/11)	69 (20/9)	58 (15/11)	73 $ (16/6)
5.1	B72-3*	Yuan - 2009[r39]	pleural, peritoneal, pericardial	100 (36/0)	69 (65/29)		

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
5.2	B72-3*	Grefte - 2007[r42]	pleural, peritoneal	100 (11/0)	42 (5/7)	100 (11/0)	0 (0/11)
5.3	B72-3*	Delahaye - 1997[r67]	unknown	98 (40/1)	77 (68/20)	98 (40/1)	0 (0/25)
5.4	B72-3*	Kuhlman - 1991[r73]	pleural	90 (18/2)	95 (19/1)	100 (20/0)	0 $ (0/20)
5.5	B72-3*	Betta - 1991[r71]	pleural, peritoneal	90 (9/1)	80 (16/4)		
6.1	HMFG-2*	Grefte - 2007[r42]	pleural, peritoneal	36 (4/7)	100 (12/0)	36 (4/7)	0 (0/11)
6.2	HMFG-2*	Linari - 1989[r74] ^	pleural, peritoneal, pericardial	9 (1/10)	71 (146/60)	9 (1/10)	0 (0/15)
6.3	HMFG-2*	Ghosh - 1987[r76]	pleural, peritoneal	17 (2/10)	63 (5/3)	17 (2/10)	0 (0/5)
6.4	HMFG-2*	Cibas - 1987[r75]	pleural, peritoneal	20 (4/16)	95 (37/2)		
7.1	cytokeratin CK5/6	Shield - 2007[r45]	pleural, peritoneal	97 (33/1)	91 (61/6)		
7.2	cytokeratin CK5/6	Saad - 2006[r49]	pleural	90 (18/2)	50 (10/10)		
7.3	cytokeratin CK5/6	Li - 2006[r48]	pleural	92 (11/1)	57 (12/9)		
8.1	D2-40	Bhalla - 2007[r40]	pleural, peritoneal	100 (10/0)	100 (10/0)	100 (10/0)	20 (4/16)
8.2	D2-40	Saad - 2006[r49]	pleural	85 (17/3)	100 (20/0)		
8.3	D2-40	Bassarova - 2006[r47]	pleural, peritoneal, pericardial	94 (30/2)	92 (230/20)	94 (30/2)	0 (0/8)
9.1	Leu-M1*	Dejmek - 1999[r63]	pleural	86 (30/5)	51 (24/23)	86 (30/5)	4 (1/23)
9.2	Leu-M1*	Motherby - 1999[r64]	pleural, peritoneal, pericardial, cul de sac	100 (14/0)	32 # (28/59)	100 (14/0)	0 $ (0/53)
9.3	Leu-M1*	Delahaye - 1997[r67]	unknown	100 (41/0)	28 (25/63)	100 (41/0)	0 (0/25)
10.1	MOC-31*	Pu - 2007[r44]	pleural, peritoneal	67 (12/6)	80 (20/5)		
10.2	MOC-31*	Hecht - 2005[r52]	pleural, peritoneal, pericardial	86 (6/1)	100 (86/0)	94 (16/1)	11 (1/8)
10.3	MOC-31*	Delahaye - 1997[r67]	unknown	88 (36/5)	76 (67/21)	88 (36/5)	0 (0/25)
11.1	TTF-1*	Saad - 2006[r49]	pleural	100 (20/0)	45 (9/11)		
11.2	TTF-1*	Afify - 2002[r56]	unknown	100 (12/0)	39 (27/43)		
11.3	TTF-1*	Hecht - 2001[r60] ^	pleural, peritoneal, pericardial	100 (14/0)	39 (37/57)		
12.1	vimentin	Dejmek - 2005[r82]	unknown	84 (43/8)	50 (60/59)		

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
12.2	vimentin	Davidson - 2001[r58]	pleural, peritoneal	75 (9/3)	67 (66/32)	75 (9/3)	11 (6/50)
12.3	vimentin	Kuhlman - 1991[r73]	pleural	95 (19/1)	80 (16/4)	85 (17/3)	15 $ (3/17)
13.1	Ca125	Dejmek - 2005[r82]	unknown	92 (12/1)	34 (24/46)		
13.2	CA125	Davidson - 2001[r58]	pleural, peritoneal	100 (12/0)	29 (28/70)	100 (12/0)	16 (9/47)
14.1	cytokeratin	Ascoli - 1995[r68 ^]	pleural, peritoneal, hydrocoele, synovial	100 (33/0)	5 (7/145)	100 (33/0)	0 $ (0/35)
14.2	cytokeratin	Kuhlman - 1991[r73]	pleural	100 (20/0)	0 (0/20)	95 (19/1)	40 $ (8/12)
15.1	cytokeratin CAM5-2	Li - 2006[r48]	pleural	100 (12/0)	10 (2/19)		
15.2	cytokeratin CAM5-2	Dejmek - 1999[r63]	pleural	97 (33/1)	2 (1/50)	97 (33/1)	0 (0/24)
16.1	desmin*	Afify - 2002[r57]	pleural, peritoneal	100 (14/0)	0 (0/56)	100 (14/0)	92 $ (22/2)
16.2	desmin*	Davidson - 2001[r58]	pleural, peritoneal	92 (11/1)	2 (2/96)	92 (11/1)	84 (47/9)
17.1	E-cadherin	Sivertsen - 2006[r50]	pleural, peritoneal	58 (14/10)	13 (7/46)		
17.2	E-cadherin	Simir - 1999[r65]	pleural	46 (12/14)	3 (1/28)	46 (12/14)	86 $ (19/3)
18.1	HBME-1	Dejmek - 2005[r82]	unknown	0 (0/7)	23 (14/47)		
18.2	HBME-1	Ascoli - 1997[r66 ^]	pleural, peritoneal, pericardial, hydrocoele	100 (47/0)	76 (95/30)		
19.1	keratin	Cibas - 1987[r75]	pleural, peritoneal	100 (20/0)	0 (0/39)		
19.2	keratin (kDalton63)	Walts - 1983[r77 ^]	unknown			83 (10/2)	0 (0/15)
20.1	mesothelin	Pu - 2007[r44]	pleural, peritoneal	44 (8/10)	24 (6/19)		
20.2	mesothelin	Donna - 1992[r70]	pleural, peritoneal	100 (12/0)	100 (12/0)		
21.1	N-cadherin	Sivertsen - 2006[r50]	pleural, peritoneal	63 (15/9)	25 (13/40)		
21.2	N-cadherin	Simir - 1999[r65]	pleural	35 (9/17)	52 (15/14)	35 (9/17)	23 $ (5/17)
22.1	p63*	Pu - 2007[r44]	pleural, peritoneal	100 (18/0)	60 (15/10)		
22.2	p63*	Saad - 2006[r49]	pleural	100 (20/0)	55 (11/9)		
23.1	thrombomodulin	Dejmek - 2005[r82]	unknown	86 (6/1)	53 (31/28)		
23.2	thrombomodulin	Ascoli - 1995[r68 ^]	pleural, peritoneal, hydrocoele, synovial	100 (33/0)	62 (94/58)	100 (33/0)	0 $ (0/35)

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
24.1	WT-1	Pu - 2007[r44]	pleural, peritoneal	100 (18/0)	100 (25/0)		
24.2	WT-1	Saad - 2006[r49]	pleural	95 (19/1)	100 (20/0)		
25	actin*	Afify - 2002[r57]	pleural, peritoneal	100 (14/0)	0 (0/56)	100 (14/0)	0 $ (0/24)
26	BMA-120	Kuhlman - 1991[r73]	pleural	85 (17/3)	90 (18/2)	80 (16/4)	20 $ (4/16)
27	CA1/2	Ghosh - 1987[r76]	pleural, peritoneal	75 (9/3)	38 (3/5)	75 (9/3)	100 (5/0)
28	CD138*	Saqi - 2005[r54 ^]	pleural, peritoneal, pericardial	92 (22/2)	44 (19/24)	92 (22/2)	0 (0/8)
29	CD44S	Afify - 2005[r51]	pleural, peritoneal	86 (12/2)	62 (38/23)	86 (12/2)	0 $ (0/28)
30	claudin1*	Kleinberg - 2007[r43]	pleural, peritoneal, pericardial	80 (20/5)	71 (212/88)		
31	claudin3*	Kleinberg - 2007[r43]	pleural, peritoneal, pericardial	100 (25/0)	59 (177/123)		
32	claudin4*	Facchetti - 2007[r41]	pleural, peritoneal	100 (23/0)	97 (60/2)	100 (23/0)	0 $ (0/12)
33	cytokeratin CK5	Bhalla - 2007[r40]	pleural, peritoneal	100 (10/0)	90 (9/1)	100 (10/0)	0 (0/20)
34	cytokeratin K903	Li - 2006[r48]	pleural	92 (11/1)	52 (11/10)		
35	Glut-1m	Shen - 2009[r37]	pleural, peritoneal			63 (22/13)	82 (31/7)
36	Glut-1p	Shen - 2009[r37]	pleural, peritoneal			83 (29/6)	63 (24/14)
37	HA	Afify - 2005[r51]	pleural, peritoneal	100 (14/0)	100 (61/0)	100 (14/0)	7 $ (2/26)
38	HEA-125*	Kuhlman - 1991[r73]	pleural	95 (19/1)	95 (19/1)	100 (20/0)	0 $ (0/20)
39	IGF-II*	Slipicevic - 2009[r38]	pleural, peritoneal, pericardial	70 (23/10)	59 (172/122)		
40	IGFBP3*	Slipicevic - 2009[r38]	pleural, peritoneal, pericardial	76 (25/8)	49 (144/150)		
41	keratin7*	Baars - 1994[r69]	pleural, peritoneal	100 (10/0)	73 (46/17)	100 (10/0)	20 $ (3/12)
42	Ki67	Schonherr - 2004[r55]	pleural, peritoneal, pericardial			25 (5/15)	100 $ (20/0)
43	myogenin*	Afify - 2002[r57]	pleural, peritoneal	100 (14/0)	0 (0/56)	100 (14/0)	0 $ (0/24)
44	myoglobin*	Afify - 2002[r57]	pleural, peritoneal	100 (14/0)	0 (0/56)	100 (14/0)	0 $ (0/24)
45	OV632	Delahaye - 1991[r72]	pleural, peritoneal	92 (22/2)	68 (21/10)	92 (22/2)	100 (20/0)
46	P-cadherin	Sivertsen - 2006[r50]	pleural, peritoneal	83 (20/4)	38 (20/33)		

Table 2.3: Reported sensitivity and specificity of immunohistochemical markers per study stratified by type marker *(Continued)*

No.	Marker	1st Author-year	Type of effusion(s)	In comparison with			
				Malignancy		Non-malignancy	
				Sens % TP/FN	Spec % (TN/FP)	Sens % TP/FN	Spec % (TN/FP)
47	p53	Davidson - 2001[r58]	pleural, peritoneal	83 (10/2)	33 (32/66)	83 (10/2)	86 (48/8)
48	podoplanin	Bhalla - 2007[r40]	pleural, peritoneal	100 (10/0)	100 (10/0)	100 (10/0)	15 (3/17)
49	Sial-Tn	Dejmek - 2005[r82]	unknown	71 (5/2)	52 (30/28)		
50	SV-40*	Simsir - 2001[r61]	pleural	100 (32/0)	0 (0/43)	100 32/0)	0 (0/25)
51	TAG-72*	Aerts - 2006[r46]	pleural	100 (14/0)	92 (11/1)	100 (14/0)	0 (0/13)
52	Tenascin-X	Yuan - 2009[r39]	pleural, peritoneal, pericardial	76 (28/9)	97 (133/4)	76 (28/9)	89 [$] (8/1)
53	WT1	Hecht - 2001[r59]	pleural, peritoneal, pericardial	100 (14/0)	77 (75/22)		
54	XIAP	Shen - 2009[r37]	pleural, peritoneal			83 (29/6)	39 (15/23)

*sensitivity and specificity were calculated as follows: the number of mesothelioma patients below the cut-off value was defined as TP and those above the cut-off value as FN whereas the number of non-mesothelioma patients below the cut-off value was defined as FN and those above the cut-off value as TN; ^(r39) for any staining pattern, a two-by-two table was constructed by combining the values of membranous and cytoplasmic staining; ^(r42, r75) membranous staining was classified as a predominant staining of the mebrame (i.e. the membranous staining was substantially greater than the cytoplasmic); [#](r64) values of other malignant diseases included two patients of chronic unspecific pleuritis; [$](r64) values of non-malignancy included one patient with lung cancer; ^(r68) the values of group IV were excluded from the two-by-two table because these were highly influenced by incorporation bias; [$](r68) 40% of the samples that contained benign reactive cells were obtained from patients with underlying malignancies; [$](r73) we considered samples with benign reactive cells as benign, although the underlying diseases were not described; ^(r81) other anti-EMA clones were also studied, but E29 and Mc5 were the most useful clones.

Chapter 4

Prognosis and prognostic factors of patients with mesothelioma: a population based study

S van der Bij
H Koffijberg
JA Burgers
P Baas
MJ van de Vijver
BAJM de Mol
KGM Moons

Published in Br J Cancer 2012: 107(1): 161-4

Abstract

Background: It is important to regularly update survival estimates of patients with malignant mesothelioma as prognosis may vary according to epidemiologic factors and diagnostic and therapeutic management.

Methods: We assessed overall (baseline) survival as well as related prognostic variables in a large cohort of 1,353 patients with a confirmed diagnosis of malignant mesothelioma between 2005-2008.

Results: About 50% of the patients were 70 years or older at diagnosis and the median latency time since start of asbestos exposure was 49 years. One year after diagnosis, 47% of the patients were alive, 20% after two years and 15% after three years. Prognostic variables independently associated with worse survival were: older age (HR=1.04 per year 95%CI [1.03-1.06]), sarcomatoid subtype (HR=2.45 95%CI [2.06-2.90]), and non-pleural localization (HR=1.67 95%CI [1.26-2.22]).

Conclusion: Survival of patients with malignant mesothelioma is still limited and depends highly on patient age, mesothelioma subtype and localization. In addition, a substantial part of the patients had a long latency time between asbestos exposure and diagnosis.

Introduction

The prognosis of patients with malignant mesothelioma is usually poor. However, a small fraction of patients is still alive two years after diagnosis. Differences in survival are associated with age at diagnosis, gender, health status and tumour and environment related factors.[1] Several studies have shown that asbestos exposure is negatively correlated with prognosis of patients with malignant mesothelioma, and therefore prognosis can vary across regions with different histories of industrial exposure to asbestos.[2;3] However, geographical differences might also reflect local approaches to diagnostic and therapeutic management. Any delay in making the diagnosis of malignant mesothelioma may have a major effect on survival estimates when survival is as short as in mesothelioma patients and can affect associations between prognostic factors and survival. Differences in prognosis resulting from differences in asbestos exposure and management of malignant mesothelioma might be expected across countries and over time. Therefore, it is important to regularly, and regionally, update survival estimates, with use of population-based studies.

The aim of this study, based on recent evidence from a large population-based cohort in The Netherlands, was threefold: 1) to update survival estimates of patients with malignant mesothelioma, 2) to identify general predictors of survival, and 3) to assess the predictive accuracy of the combined prognostic factors for prolonged survival.

Methods

Patients

This study involves retrospective analyses of an existing registry comprising 1,353 patients with malignant mesothelioma who applied to the Dutch Institute for Asbestos Victims and entered the process for getting financial compensation between 2005 and 2008.[Baas et al., 2006].

After application, the standard procedure is that the patients or relatives were visited at home by a qualified representative of the Dutch asbestos institute who further explained the application procedure and compensation scheme. Patients who decided to participate in this compensation scheme had to give written informed consent for the use of their clinical data and data regarding social status, occupational circumstances and income by the institute for assessing their case and for internal and external analyses and reporting. For such linking and use of data, the most strict rules and potential sanctions are applicable regarding confidentiality and anonymization. Accordingly, in the present analyses data were also completely anonymized. The diagnosis of mesothelioma was confirmed by pathologists from the Dutch National Mesothelioma Panel (NMP). When pathological material was not available or insufficient for a confirmed diagnosis by the NMP (N=62), a final diagnosis was reached by three independent pulmonologists of the Mesothelioma Group of the Dutch Thoracic Society (DTS).[4]

Survival (outcome)

Survival was measured from the date of clinical diagnosis till death or censoring. The date of diagnosis was defined as the date on which malignant mesothelioma was diagnosed by the local hospital.

Prognostic variables

The following prognostic variables were studied: gender, age at diagnosis, year of diagnosis, pathologic morphologic subtype (epithelial, sarcomatoid and biphasic type), tumour location (pleural, peritoneal or other), and various variables associated with asbestos exposure, i.e. duration of asbestos exposure, latency time (defined as the time elapsed between first asbestos exposure and diagnosis) and direct exposure (yes, no).

Analysis

For the first aim, assessing overall survival, Kaplan-Meier analyses were used and survival curves were plotted. To put results into perspective, survival probabilities of our cohort were compared with the overall survival in the general Dutch population after adjustment for gender and age.

For the second aim, associations between possible prognostic variables and survival were estimated using Cox proportional-hazards regression. Missing values were imputed with multiple imputation, conforming with current guidelines, since missing values occurred on various predictor variables (see Appendix). [5] Bootstrapping was used to correct for overfitting.[6;7]

Finally, for our third aim, we estimated the predictive accuracy of all prognostic variables combined using both discrimination and calibration statistics. The discrimination was tested by Harrell's c-statistic for censored data, and corrected for overfitting.[7] The calibration was performed based on survival after one year. Predicted one-year survival probabilities were calculated according to different prognostic factors.

All analyses were performed with SAS enterprise guide 4.3 (SAS Institute Inc., Cary, NC, USA) and R version 2.10.1 (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Patients

The baseline characteristics of the 1,353 included patients are described in Table 1 (left column). In our cohort, the mean age at time of diagnosis was 69 years and the majority of patients was male (91%). In almost all patients (96%) the tumour was located in the pleura. Epithelial morphology was the most frequent mesothelioma type (78%). In 78% of the patients, a history of direct exposure to asbestos was identified. About half of the patients had an asbestos exposure duration ≥ 20 years. The latency time since first exposure ranged from 19 to 78 years with a median of 49 years.

Table 1: Baseline characteristics and their unadjusted and adjusted effect on survival in patients diagnosed with malignant mesothelioma.

Variables	Patients (n=1,353) (values are numbers (%) unless stated otherwise)	Crude HR (95%CI)	Adjusted HR (95%CI) [c]
***Age** (in years)*			
<60	209 (15.4)	1.00 (reference)	1.00 (reference)
60 - <70	499 (36.9)	1.94 (1.58-2.37)**	1.82 (1.46-2.26)**
70 - <80	541 (40.0)	2.83 (2.32-3.46)**	2.47 (1.93-3.17)**
≥ 80	104 (7.7)	3.83 (2.93-5.01)**	3.38 (2.45-4.65)**
age (as continuous variable, mean ± sd)	69 (±8)	1.05 (1.04-1.06)**	1.04 (1.03-1.06)**
median age (min-max)	69 (39-95)		
Gender			
Female	120 (8.9)	1.00 (reference)	1.00 (reference)
Male	1,233 (91.1)	1.40 (1.13-1.73)**	1.16 (0.93-1.46)
Tumour location			
Pleural	1,296 (95.8)	1.00 (reference)	1.00 (reference)
non-pleural/ peritoneal [a]	57 (4.2)	1.39 (1.05-1.85)**	1.67 (1.26-2.22)**
Mesothelioma morphology (pathologic subtype)			
Epithelial	1,049 (77.5)	1.00 (reference)	1.00 (reference)
Sarcomatoid	209 (15.4)	2.50 (2.13-2.93)**	2.45 (2.06-2.90)**
Mixed	95 (7.0)	1.59 (1.27-1.99)**	1.65 (1.32-2.06)**
Asbestos exposure			
Duration of asbestos (in years)			
<5 years	143 (10.6)	1.00 (reference)	1.00 (reference)
5 - <10 years	164 (12.1)	1.12 (0.87-1.44)	1.18 (0.92-1.52)
10 - <20 years	423 (31.3)	1.21 (0.98-1.49)	1.21 (0.98-1.49)
20 - <30 years	321 (23.7)	1.25 (1.00-1.58)	1.19 (0.95-1.48)
≥ 30 years	302 (22.3)	1.59 (1.27-1.98)**	1.28 (1.02-1.61)*
duration (as continuous variable, mean ± sd)	20 (±12)	1.01 (1.01-1.02)**	1.00 (1.00-1.01)
median duration (min-max)	19 (1-66)		
Latency time (in years)			
<40 years	268 (19.8)	1.00 (reference)	1.00 (reference)
40 - <50 years	487 (36.0)	1.24 (1.03-1.48)**	0.94 (0.77-1.15)
≥ 50 years	598 (44.2)	1.98 (1.66-2.36)**	1.10 (0.86-1.40)
latency (as continuous variable, mean ± sd)	48 (±9)	1.03 (1.03-1.04)**	1.00 (0.99-1.01)
median latency (min-max)	49 (19-78)		
Direct exposure of asbestos [b]	1,052 (77.8)	1.14 (0.99-1.32)	1.01 (0.87-1.17)

HR = hazard rates; CI= confidence limit; [a] Including one patient with pericardial mesothelioma, all other patients had peritoneal mesothelioma; [b] in comparison to second-hand exposure and no distinct asbestos exposure; [c] hazard rates after shrinking, results of the multivariable model are based on the inclusion of the continuous variables as linear terms, the model was refitted for the estimation of the HRs of the continuous variables as categorical variables; *significant at a p-value of 0.05, ** significant at a p-value of 0.01 (the overall p-value was also checked for categorical variables with more than two categories and was significant in the multivariable model for age and pathologic subtype (p value of <0.01)). Overall, the most significant predictor in the multivariable model was pathologic subtype.

Overall Survival

Figure 1 shows the overall survival curve from the time of clinical diagnosis. Median survival was 333 days (95%CI:309-368); 47% of the patients survived longer than one year and 20% longer than two years. Less than 15% of the patients were alive three years after diagnosis in contrast to 90% of individuals with similar age and gender distribution in the general Dutch population.

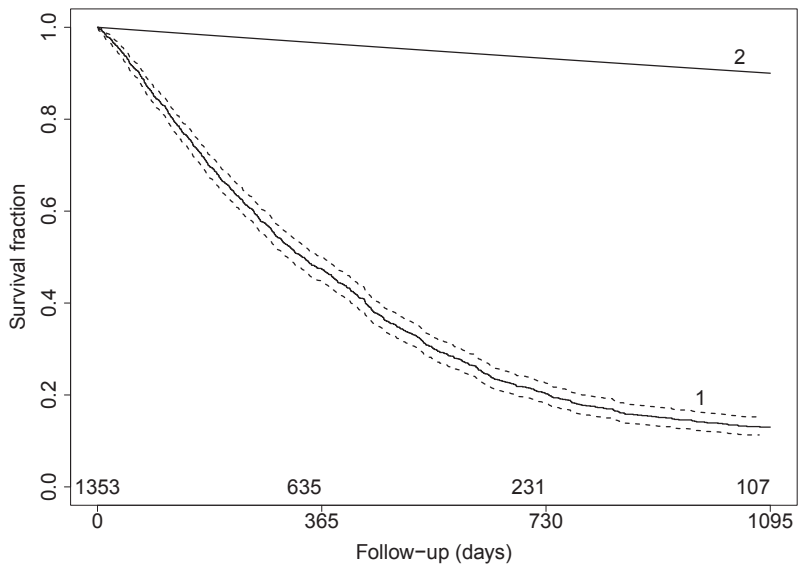


Figure 1: Kaplan Meier survival curve showing the overall survival and 95%CI from the time of the diagnosis of mesothelioma for the entire study cohort (1) and of the general Dutch population (2). The 95%CI is presented by the broken line. The Dutch population was adjusted (i.e. standardized) to the age and gender distribution of the study cohort. The number of study patients at risk is indicated at the bottom of the plot (above the x-axis).

Prognostic factors and predictive accuracy

In a univariable analysis, all variables, except direct exposure, were significantly associated with survival at a significance level of 0.05 (Table 1, middle column). In the multivariable model, only age, morphology, localization of malignant mesothelioma had a significant independent association with survival (Table 1, last column). Hence, worse survival was independently associated with older age, sarcomatoid subtype, or non-pleural localization.
Table 2 shows the predicted one-year survival probabilities stratified by tumour location, pathologic subtype and age based on our multivariable model. The one-year survival given a diagnosis of pleural malignant mesothelioma of epithelial subtype was estimated to be 77% for a patient of 50 years and 38% for a patient of 80 years. Conversely, the one-year survival given a diagnosis of pleural malignant mesothelioma of sarcomatoid subtype was estimated to be 53% for a patient of 50 years and 9% for a patient of 80 years. These estimated survival probabilities are much lower than those in the general Dutch population, where a man aged 50 or aged 80 has a one year survival probability of 99.7% and 92.5%

respectively. The multivariable model showed a c-statistic of 0.66 (95%CI: 0.64-0.68) and very good calibration (Hosmer-Lemeshow X^2=7.63, p-value=0.57).

Table 2: Predicted one-year survival from the time of the diagnosis of mesothelioma stratified by tumour location, pathologic subtype and age [a]

	predicted one year survival of patients with pleural mesothelioma (%)	predicted one year survival of patients with peritoneal mesothelioma (%)
for epithelial subtype:		
age: 50 years	77	65
60 years	67	51
70 years	53	35
80 years	38	20
for sarcomatoid subtype:		
age: 50 years	53	34
60 years	37	19
70 years	22	8
80 years	9	2

[a] Results were based on the average values over the other covariates.

Discussion

To date, the survival of malignant mesothelioma patients remains poor. After one year, only 47% of the patients were still alive. Predictors strongly associated with survival were patient age, mesothelioma localisation and subtype. These results are consistent with other population-based studies.[8-11] This study showed that the discriminative ability of these general predictors was moderate and the calibration was good.

Our observed survival was only marginally higher than in two older Dutch studies, in which survival among patients diagnosed with malignant mesothelioma between 1970-1994 and 1987-1989 was studied.[12;13] In these studies the probability of one-year survival was about 42% suggesting that survival has not improved substantially over the years. Lack of improvement was also observed by a recent Italian and American study.[9;11] However, if the mix of patients has changed over the years due to, for example, improved diagnosis in patients with suspected mesothelioma, then direct comparisons between older studies and our study are hard to make. Moreover, in the Dutch study of van Gelder et al.[13], in which patients diagnosed between 1987-1989 were studied, 42% of the patients were 65 years or younger, whereas in our study only 30% of the patients were younger than 65 years (data not shown).
The higher age in our cohort likely relates to currently longer latency times between asbestos exposure and diagnosis. Our results showed an average latency time of 49 years between initial asbestos exposure and diagnosis.

The prognostic value of patient age, malignant mesothelioma subtype and localization can assist in the selection of patients more likely to benefit from intensive treatment modalities, especially for patient selection in future therapeutic randomized trials. However, in the current study not all potentially relevant predictors were available that might contribute to the discrimination of survival among malignant mesothelioma patients. For example, there is some evidence that patients' general well-being and weight loss are important prognostic factors in patients with malignant mesothelioma.[14-16] Therefore, we expect that predictive accuracy might improve when these predictors would also be taken into account. We did not observe a significant association between characteristics of asbestos exposure and survival. However, in our data set asbestos exposure was mainly based on self-reporting, which could mean that exposure estimates are of lower quality than the other predictors considered.

Recently, more treatment options have become available for patients with malignant mesothelioma. Although these may benefit selected patients, their results are still far from satisfactory for the majority of the patients.[17] In our study, patients received treatment according to latest insights suggesting that, in general, the impact of treatment is still limited. To improve the effect of treatment, an early diagnosis of malignant mesothelioma is of great value. This may hold in particular for patients with peritoneal mesothelioma, as the observed difference in survival between peritoneal mesothelioma and pleural mesothelioma may be explained by a delayed diagnosis of peritoneal mesothelioma due to the complexity of the disease.[18]

In conclusion, we showed that overall survival in patients with malignant mesothelioma remains poor and depends highly on patient age, malignant mesothelioma subtype and localization. Additionally, we found that half of the patients are 70 years or older and a substantial part of the patients has a long latency time since asbestos exposure. A trend towards longer latency times may have profound implications for future lawsuits and reimbursements as in several countries financial compensation depends (partially) on latency times.[4] Furthermore, the future prevalence of mesothelioma might still remain high as a result of these long latency times.

Acknowledgements

This study was funded by the Institute for Asbestos Victims, the Netherlands

Disclaimer

B.A.J.M.M is a board member of the Institute for Asbestos Victims. M.J.V is chairman of the Dutch National Mesothelioma Panel. J.A.B. is chairman of the Mesothelioma Group of the Dutch Thoracic Society. Conclusions of this paper reflect the opinions of the authors and do not represent any determination or policy of the Institute for Asbestos Victims.

Reference List

1. Burgers JA, Damhuis RA. Prognostic factors in malignant mesothelioma. Lung Cancer 2004; 45 Suppl 1:S49-S54.
2. Christensen BC, Godleski JJ, Roelofs CR, Longacker JL, Bueno R, Sugarbaker DJ et al. Asbestos burden predicts survival in pleural mesothelioma. Environ Health Perspect 2008; 116:723-726.
3. Spirtas R, Connelly RR, Tucker MA. Survival patterns for malignant mesothelioma: the SEER experience. Int J Cancer 1988; 41:525-530.
4. Baas P, van 't Hullenaar N, Wagenaar J, Kaajan JP, Koolen M, Schrijver M et al. Occupational asbestos exposure: how to deal with suspected mesothelioma cases--the Dutch approach. Ann Oncol 2006; 17:848-852.
5. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006; 59:1087-1091.
6. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ 2009; 338:b604.
7. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001.
8. Mirabelli D, Roberti S, Gangemi M, Rosato R, Ricceri F, Merler E et al. Survival of peritoneal malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:194-200.
9. Montanaro F, Rosato R, Gangemi M, Roberti S, Ricceri F, Merler E et al. Survival of pleural malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:201-207.
10. Chapman A, Mulrennan S, Ladd B, Muers MF. Population based epidemiology and prognosis of mesothelioma in Leeds, UK. Thorax 2008; 63:435-439.
11. Milano MT, Zhang H. Malignant pleural mesothelioma: a population-based study of survival. J Thorac Oncol 2010; 5:1841-1848.
12. Janssen-Heijnen ML, Damhuis RA, Klinkhamer PJ, Schipper RM, Coebergh JW. Increased but low incidence and poor survival of malignant mesothelioma in the southeastern part of The Netherlands since 1970: a population-based study. Eur J Cancer Prev 1999; 8:311-314.
13. van Gelder T, Damhuis RA, Hoogsteden HC. Prognostic factors and survival in malignant pleural mesothelioma. Eur Respir J 1994; 7:1035-1038.
14. Ak G, Metintas S, Metintas M, Yildirim H, Erginel S, Kurt E et al. Prognostic factors according to the treatment schedule in malignant pleural mesothelioma. J Thorac Oncol 2009; 4:1425-1430.
15. Nowak AK, Francis RJ, Phillips MJ, Millward MJ, van der Schaaf AA, Boucek J et al. A novel prognostic model for malignant mesothelioma incorporating quantitative FDG-PET imaging with clinical parameters. Clin Cancer Res 2010; 16:2409-2417.
16. Edwards JG, Abrams KR, Leverment JN, Spyt TJ, Waller DA, O'Byrne KJ. Prognostic factors for malignant mesothelioma in 142 patients: validation of CALGB and EORTC prognostic scoring systems. Thorax 2000; 55:731-735.
17. Mott FE. Mesothelioma: a review. Ochsner J 2012; 12:70-79.
18. Manzini VP, Recchia L, Cafferata M, Porta C, Siena S, Giannetta L et al. Malignant peritoneal mesothelioma: a multicenter study on 81 cases. Ann Oncol 2010; 21:348-353.

Appendix

Appendix: Distribution of missing values among available variables (including the outcome).

Variables	Missing values	Complete cases (patients with all variables completely observed)	Patients with at least one missing value	P-value [b]
	number (%)	n=686 (51%)	n=667 (49%)	
Age *(in years) mean (± sd)*	156 (12)	69 (±8)	69 (±9)	0.294
Male gender	1 (0)			<.001
Female		18 (3)	102 (15)	
Male		668 (97)	564 (85)	
Tumour location	90 (7)			0.001
Pleural		667 (97)	539 (93)	
non-pleural/ peritoneal		19 (3)	38 (7)	
Mesothelioma morphology *(pathologic subtype)*	72 (5)			0.880
Epithelial		520 (76)	457 (77)	
Sarcomatoid		113 (16)	96 (16)	
Mixed		53 (8)	42 (7)	
Asbestos exposure				
mean duration of exposure in years (± sd)	314 (23)	20 (±13)	20 (±14)	0.545
mean latency time in years (± sd)	360 (27)	48 (±9)	48 (±9)	0.830
direct exposure of asbestos	523 (39)	574 (84)	67 (47)	<.001
Survival *mean time in days (± sd)*	156 (12) [a]	444 (±395)	424 (±385)	0.403

[a] missings are due to missing data on the actual date of diagnosis ; [b] Significant (Underlined) P-values indicate that missing data were clearly not missing completely at random (MCAR) but related to the observed variables. Hence, patients with missing values were no random but rather a selective subset of the total cohort. Simply excluding this selective subset of subjects by performing a so-called 'complete subject analysis' would thus lead to biased results. Therefore, missing values were imputed with multiple regression techniques.

Part 2

Asbestos-related lung cancer

Chapter 5

Lung cancer risk at low cumulative asbestos exposure: meta-regression of the exposure-response relationship

S van der Bij
H Koffijberg
V Lenters
L Portengen
KGM Moons
D Heederik
RCH Vermeulen

Submitted

Abstract

Background: Existing estimated lung cancer risks per unit of asbestos exposure are mainly based on, and applicable to, high exposure levels. To assess the risk at low cumulative asbestos exposure we provide new evidence by fitting flexible meta-regression models, a notably new and more robust method.

Methods: Studies were selected if lung cancer risk per cumulative asbestos exposure in at least two exposure categories was reported. From these studies (n=19), we extracted 104 risk estimates over a cumulative exposure range of 0.11 to 4,710 f-y/ml. We fitted linear and natural spline meta-regression models to these risk estimates. A natural spline allows risks to vary non-linearly with exposure, such that estimates at low exposure are less affected by estimates in the upper exposure categories. Associated relative risks (RRs) were calculated for several low cumulative asbestos exposures.

Results: A natural spline model fitted our data best. With this model the relative lung cancer risk for cumulative exposure levels of 4 f-y/ml, and 40 f-y/ml was estimated between 1.013 and 1.027, and 1.13 and 1.30, respectively. After stratification by fibre type, a non-significant 3 to 4-fold difference in RRs between chrysotile and amphibole fibres was found for exposures below 40 f-y/ml. Fibre type-specific risk estimates were strongly influenced by a few studies.

Conclusions: The natural spline regression model indicates that at lower asbestos exposure levels, the increase in RR of lung cancer due to asbestos exposure may be larger than expected from previous meta-analyses. Observed potency differences between different fibre types are lower than the generally held consensus. Low exposed industry or population-based cohorts with quantitative estimates of asbestos exposure are required to substantiate the risk estimates at low exposure levels from our new, flexible meta-regression.

Introduction

It is widely accepted that exposure to asbestos is related to an excess risk of lung cancer.[1] However, studies exploring the exposure-response relationship have shown a large variability in excess risk per unit of exposure. Berman and Crump showed in a meta-analysis that such variations might be related to different fibre size distributions and fibre type [2]. Within fibre type, relatively longer fibres were associated with a higher increased lung cancer risk compared to shorter fibres. Moreover, chrysotile was estimated to be less potent than amphiboles by a factor ranging between 6 and 60. The meta-analysis of Hodgson and Darnton had comparable findings with respect to potency differences between chrysotile and amphiboles.[3] In a recent meta-analysis by Lenters et al. it was shown that variations in risk estimates of lung cancer might additionally be explained by differences in quality aspects of the applied exposure assessments methodology besides fibre type.[4]

In all previous meta-analyses[2;4;5] except for the one of Hodgson and Darnton[3], a fixed, average excess risk per fibre year (expressed as the potency, i.e. the so-called K_L value, of asbestos for causing lung cancer) was estimated by combining the K_L-values obtained for each study. However, most of the lung cancer studies included in the meta-analyses were not very recent and notably involved heavily exposed individuals. Currently, certainly in the Western world, it is unlikely that individuals are exposed to levels previously generally studied, because handling of asbestos declined gradually after the 1970s and dropped severely in the 1990s due to directives on protecting workers exposed to asbestos.[6] As a consequence, current interest lies in estimating excess risk accurately at relatively low exposures.
Previously, linear extrapolation has been applied to estimate risks at low exposure levels. However, such extrapolation is heavily dependent on estimates at high exposures, rendering extrapolated risk estimates at low exposure uncertain. For example, the population-based study of Gustavsson et al. found a significant excess risk of lung cancer at low levels of cumulative asbestos exposure, which was much higher than could be expected by simple linear extrapolation from cohorts with higher exposures.[7]

To accurately derive acceptable exposure limits and underpin compensation claims, better evidence is needed about the asbestos-related risk of lung cancer at low exposures. We provide new evidence by fitting non-linear meta-regression models to existing data, which is notably new in meta-analyses. The flexibility of these models ensures that the exposure-response relationship can vary with exposure levels and is less affected by estimates in the upper exposure categories.[8] Moreover, the advantage of our method is that we combine all existing risk estimates at low exposures and obviate the need to extrapolate below the study-specific exposure range. Hence, our method provides a more robust estimate of exposure-specific lung cancer risks than previous meta-analyses. In addition, we stratified our results by fibre type to explore a potential potency difference between chrysotile and amphibole fibres at relatively low cumulative exposure levels.

Methods

Identification of included studies

The same selection criteria were applied as in the meta-analysis by Lenters et al.[4] Briefly, occupational studies from MEDLINE and EMBASE were selected if lung cancer risk per cumulative exposure in at least two exposure categories was reported. Furthermore, the cumulative exposure needed to be reducible to units of total number of fibre years (f-y/ml), which is defined as the product of the concentration of asbestos fibres per millilitre of air measured by phase contrast microscopy (PCM), and the duration of exposure in years. PCM measures fibres that are longer than 5 µm, thicker than approximately 0.25 µm, and with an aspect (length-to-width) ratio >3. Studies with only one exposure category were excluded because no study-specific exposure-response relationship could be derived. The selection criteria resulted in 18 industry-based cohort studies, including one nested case-control, and one general population-based case-control study (see table 1 for details).

Extraction of data from the incorporated studies

Information about the study design, study characteristics and exposure categories were extracted from each study. To obtain risk estimates for the 15 studies with standardized mortality ratios (SMRs), observed and expected lung cancer cases were extracted for each exposure category. The relative risks (RRs) and their confidence intervals (CIs), size of the study population and number of lung cancer cases were extracted for each exposure category with lung cancer occurrence among the two cohort studies with an assigned reference group. For the two case-control studies, the odds ratios (ORs) and their confidence intervals, and the number of lung cancer cases and controls were included. The adjusted ORs and corresponding CIs for the study of Gustavsson et al.[7] were obtained via direct communication with the authors. For the purpose of this meta-analysis, all measures of association, i.e. ORs, RRs and SMRs, were considered estimates of the RR of asbestos exposure and lung cancer occurrence.

To assign a specific point estimate of cumulative exposure to the extracted risk estimates, we used the mean of the exposure category, when described in the original publication. If not described, the midpoint of the range of the exposure categories was used. For open-ended, uppermost exposure categories, the midpoint was calculated as 5/3 times the lower bound of those categories (as proposed by the asbestos advisory committee of the Unites States environmental protection agency in 2008). For example, the midpoint estimate for an open–ended category of >100 fibre years was calculated as 5/3 * 100 = 167. For additional details on data extraction, we refer to Lenters et al.[4]

Modelling the exposure–response relation

We hereby expanded on the study of Lenters et al. in which they investigated the role of quality of the asbestos exposure assessment to potentially explain heterogeneity in linear

exposure-response slope estimates.[4] They showed that the linear exposure-response slope estimates can be influenced by measurement error. Moreover, linear extrapolations to lower exposures based on these estimates likely yields a large uncertainty as they did not focus on the actual shape of the exposure-response curve. To improve estimates in the low exposure range, we assessed the shape of the exposure-response curve by fitting non-linear meta-regression models to all available data estimates.

From the 19 studies, we extracted 104 risk estimates (i.e. study points of the RR for lung cancer at a given exposure level) over a cumulative exposure range of 0.11 to 4,710 f-y/ml. To accurately estimate associations in the lower exposure range based on all available data points, we used a previously developed macro for applying linear and non-linear regression models to the reported risk estimates.[8] In this macro the natural logarithm (LN) of the reported risk estimates was inversely weighted by their variance.[9] As risk estimates (ORs and RRs) within a single study are correlated, the variance of the risks were corrected by estimating the covariance between different risk estimates using the method of Greenland.[10] For studies reporting SMRs, no covariance was estimated as it can be assumed that the independence assumption does hold for SMRs since the total population is used as the reference group instead of a subsample.
The regression models applied consisted of full linear models (model type 1), and natural splines with prespecified knots at the 20th, 50th, and 80th percentiles (model type 2). The natural spline is a flexible model and allows risks to vary non-linearly with exposure, such that estimates at low exposure are less affected by estimates in the upper exposure categories.[11] The two model types were fitted with (option A) and without (option B) an intercept where model A assumes a difference in background rate of lung cancer between exposed and non-exposed individuals and model B assumes no difference. A model with intercept has been used in previous studies to account for potential differences in background risk.[2;4;5;12] However, if an intercept above RR=1 is due to measurement error instead of differences in background risk it is more appropriate to model the exposure-response relationship without intercept.[13] To accommodate potential between-study heterogeneity, the regression models allowed for random study-specific intercepts and exposure effects:[9]

$LN\ RR = ß_0 + ß_1 * exposure + \sigma_{u0}^2 + \sigma_{u1}^2 + \sigma_{e0}^2$ (model option A);
$LN\ RR = ß_1 * exposure + \sigma_{u1}^2 + \sigma_{e0}^2$ (model option B);
where $ß_0$ is the common intercept across studies, $ß_1$ is the common slope associated with exposure across studies, σ_{u0}^2 is the estimated variance of the intercept between studies, σ_{u1}^2 is the estimated variance of the slope between studies and σ_{e0}^2 is the variance of the individual risk estimates. (For the spline models an additional spline variable was estimated by using third order polynomials to fit a non-linear slope [11])

As a result, an additional component of the variance explaining the between-study heterogeneity was considered in weighting each observation.[14] Models were fitted using

maximum likelihood (ML) estimation and goodness of fit was assessed with the deviance (-2 log likelihood) criterion. For accurate estimation of the parameters, models were refitted using restricted maximum likelihood (REML). A variance components structure was used to compute the between-study variances for option A.

The results on the LN scale were retransformed to the 'normal' scale to identify the variation in RR as a function of exposure. We calculated the RRs and their CIs for low cumulative exposure levels of 4 and 40 f-y/ml. These levels were selected because occupational exposure standards have been endorsed from levels of 2 f/ml to 0.1 f/ml over an eight-hour time weighted average in the past decades.[15;16] Over a working life exposure of 40 years, we expect the cumulative exposure levels of workers over the last decades to be somewhere between 4 and 40 f-y/ml. For models with an intercept (option A), the predicted RR at zero exposure may not be equal to 1. To relate the estimated risk at a specific exposure level to an RR of 1 at zero exposure, models were refitted to the data points from which the common predicted intercept was subtracted. Results were stratified by fibre type (i.e. chrysotile, amphibole or mixed). For comparison, RRs were also calculated based on estimates from previous published meta-analyses.[2-5]

Sensitivity analyses

The sensitivity of the predicted risk to the inclusion of specific studies was assessed with a 'jackknife' analysis, in which studies are excluded one by one.[17] The sensitivity of the predicted risk at low exposure to the inclusion of risk estimates corresponding to high exposures (> 100 f-y/ml) was assessed by fitting models excluding these data. In addition, results were stratified to studies that included a latency in their estimates between exposure and lung cancer and studies that included no latency.

Software

All analyses were performed with SAS version 9.2 (SAS Institute Inc., Cary, NC, USA). Significance tests between fibre type-specific estimates were assessed with use of simulating the fibre type-specific risk distribution in R version 2.10.1 (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Study characteristics

Study characteristics of the 19 studies that were included in the meta-regression are listed in Table 1.

Table 1 Characteristics of the included studies

Study	Author	Study design	Type of risk estimates	Fibre type	lowest – highest exposure category (f-y/ml)	Lagged CE
1. Quebec mines and mills	Liddell et al. 1997 [27]	cohort	SMR	chrysotile	4.71 – 4710	CE to age 55
2. Italian mine and mill	Pira et al. 2009 [28]	cohort	SMR	chrysotile	50 - 666.7	CE
3. Connecticut plant	McDonald et al. 1984 [29]	cohort	SMR	chrysotile	15 - 400	CE
4. South Carolina textile plant	Hein et al. 2007 [30]	cohort	SMR	chrysotile	0.75 – 200	CE10
5. North Carolina textile plants	Loomis et al. 2009 [31]	cohort (internal reference)	RR adjusted [a]	chrysotile	5.69 – 408.3 [b]	CE10
6. Wittenoom, Australia mine	Berry et al. 2004 [32]	cohort	SMR	amphibole (crocidolite)	0.11 - 219.9	CE
7. Patterson, New Jersey insulation factory	Seidman et al. 1986 [33]	cohort	SMR	amphibole (amosite)	3 - 416.7	CE
8. Tyler, Texas, insulation manufacture	Levin et al. 1998 [34]	cohort	SMR	amphibole (amosite)	11.25 - 375.00	CE
9. Libby, Montana, Vermiculite mines and mills	Sullivan 2007 [35]	cohort	SMR	amphibole (amosite)	2.25 - 167	CE10
10. British friction products factory	Berry and Newhouse 1983 [36]	nested case-control, (matched design; internal reference)	OR	amphibole (tremolite)	29.5 – 228 [b]	CE
11. Ontario cement plant	Finkelstein 1984 [37]	cohort	SMR	mixed	15 - 250	CE
12. New Orleans cement plants	Hughes et al. 1987 [38]	cohort	SMR	mixed	4.2 - 256.2	CE10
13. Swedish cement plant	Albin et al. 1990 [39]	cohort (external reference)	RR adjusted [a]	mixed	3.1 - 88.2	CE
14. Belgium cement plant	Laquet et al. 1980 [40]	cohort	SMR	mixed	25 – 1200 [c]	CE
15. U.S. factory retirees	Enterline et al. 1987 [41]	cohort	SMR	mixed	186 – 2928	CE
16. U.S. insulation workers	Selikoff and Seidman 1991 [42]	cohort	SMR	mixed	37.5 - 375	CE10
17. Pennsylvania textiles plant	McDonald et al. [43]	cohort	SMR	mixed	15 - 330	CE10
18. Rochdale, England textile plant	Peto et al. 1985 [44]	cohort study	SMR	mixed	5.92 - 256.57	CE5
19. Stockholm County population	Gustavsson et al. 2002 [7]	case-control (matched design)	OR adjusted [a]	mixed	0.56 - 8.80	CE

[a] In the North Carolina textile plants study: results were adjusted for age, gender, race, calendar year and birth cohort; In the Swedish cement plant study: results were adjusted for age and calendar year; In the Stockholm County population study: results were adjusted for age, inclusion year, residential radon, environmental nitrogen dioxide, diesel exhaust and combustion products. [b] The lowest exposure category (corresponding to 0.383 f-y/ml in the North Carolina textile plant and 4.5 f-y/ml in the British friction factory) for which no risk estimate could be calculated was excluded from the analyses because it was used as the reference category; [c] The highest exposure category (corresponding to 2400 f-y/ml) for which no risk estimate could be calculated was excluded from the analyses because of lack of observed cases. Lagged cumulative exposure (CE) indicates whether exposures in the CE(x) years previous to follow-up were discarded.More details on the studies and their risk estimates can be found in the Supplemental Material of the study of Lenters et al.[4]

The 104 risk estimates extracted from these 19 studies and corresponding CIs are shown in Figure 1. From this figure it is apparent that the risk estimates vary substantially even at the lower end of the cumulative exposure range. Of all risk estimates, 38 (37%) were assessed at a cumulative exposure level of 40 f-y/ml or less, and 10 (10%) at a cumulative exposure level of 4 f-y/ml or less.

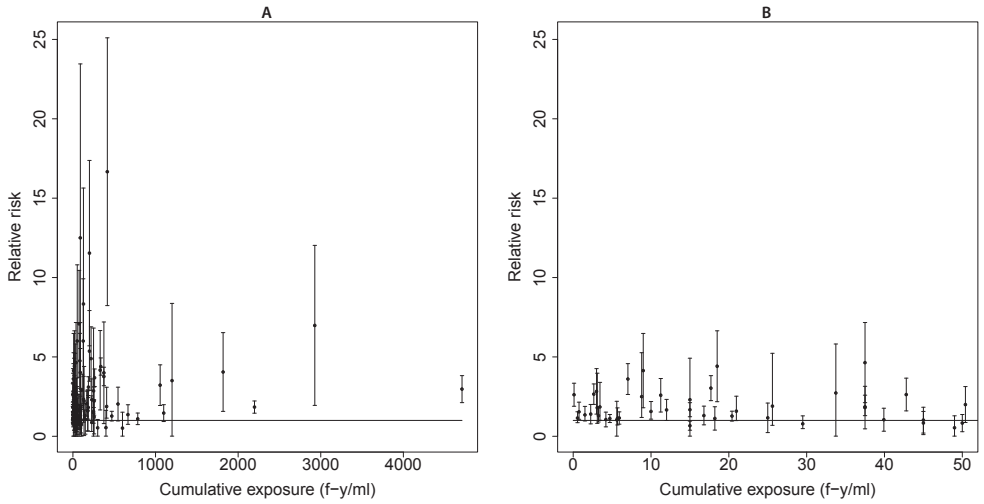


Figure 1: Scatter plot of the relative risk estimates and their 95% confidence intervals extracted from the 19 studies included in the meta-regression (A: full range of cumulative exposures and B: lower range of cumulative exposure < 50 f-y/ml)

Predictions

Table 2 shows the predicted risks based on the different exposure-response relationships.

Table 2 Comparison of predicted risk at different exposure levels

	Deviance (df) [a]	Intercept [b] (95% CI)	RR 4 f-y/ml [b] (95% CI)	RR 40 f-y/ml [b] (95% CI)
Models*				
1A. linear model	111.7 (100)	1.580 (1.243-2.008)	1.592 (1.252-2.023)	1.701 (1.338-2.164)
corrected for intercept		1.000 (0.787-1.271)	1.007 (0.793-1.280)	1.077 (0.847-1.370)
1B. linear model without intercept	806.9 (102)	1.000 (1.000-1.000)	1.017 (1.009-1.024)	1.182 (1.096-1.274)
2A. natural spline	105.6 (99)	1.483 (1.157-1.900)	1.502 (1.173-1.922)	1.680 (1.317-2.142)
corrected for intercept		1.000 (0.780-1.281)	1.013 (0.791-1.296)	1.133 (0.888-1.444)
2B. natural spline without intercept	703.6 (101)	1.000 (1.000-1.000)	1.027 (1.020-1.034)	1.301 (1.215-1.392)

RR=relative risk; df=degrees of freedom calculated as the number of data points minus the number of coefficients estimated; [a] fitted using ML estimation; [b] fitted using REML estimation; *The deviance of the empty and intercept only model was 3433.9 and 309.9 respectively.

In all models, inclusion of cumulative exposure as an explanatory variable significantly reduced model deviance. Compared to the linear model, a significantly better fit was observed for the model including a natural spline (for the explanatory variable) when a random intercept and slope was fitted (model 2A: deviance = 105.6, model 1A: deviance = 111.7; χ^2 test (1df), p = 0.01). The natural spline suggested a nearly linear increase in the relative lung cancer risk at low levels as a function of exposure (Figure 2). The slope slightly decreased after exposure of 150 f-y/ml. Based on this model, the relative lung cancer risk for 4 and 40 f-y/ml was estimated to be 1.502 (95% CI: 1.173-1.922) and 1.680 (95% CI: 1.317-2.142), respectively. After correction for the common estimated intercept, these RRs were (RR=1.013 for 4 f-y/ml and RR=1.133 for 40 f-y/ml). Similarly, when fitting regression models without intercept, a significantly better fit was observed for the spline model over the linear model (χ^2 test (1df), p = <0.001). The RR that we predicted based on the natural spline model (model 2B) was 1.027 (95% CI: 1.020-1.034) for 4 f-y/ml and 1.301 (95% CI: 1.215-1.392) for 40 f-y/ml cumulative exposure.

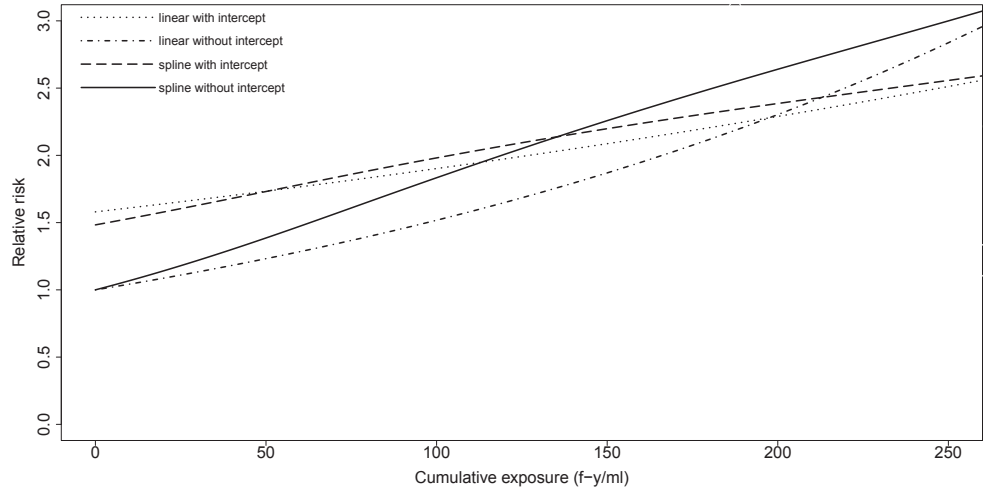


Figure 2: Predicted exposure-response relationship based on a linear and spline regression models fitted with and without intercept, shown over an exposure range of 0-250 f-y/ml (results were retransformed to the 'normal' scale)

Sensitivity analyses

A jackknife analysis, leaving one study at a time, was applied to the natural spline with intercept model (model 2A). Exclusion of the Ontario study (#11) resulted in the highest slope estimates, whereas the Pennsylvania study (#17) resulted in the lowest. However, their influence on the predicted risks was negligible after correction for the intercept (data not shown). After correction for the intercept the predicted risks increased most, and considerably, upon exclusion of the Quebec study (#1) from the analysis (RR=1.019 for 4 f-y/ml and RR=1.211 for 40 f-y/ml), and decreased most, but only slightly, upon exclusion of the South Carolina study (#4) (RR=1.010 for 4 f-y/ml and RR=1.103 for 40 f-y/ml).

When models were fitted on exclusively risk estimates corresponding to exposures of 100 f-y/ml or less, a non-significant better fit was observed for the spline model as compared to a linear model (Appendix Table A1). Based on this sensitivity analyses, predicted risk ranged from RR=1.012 for 4 f-y/ml to RR=1.152 for 40 f-y/ml, which are comparable to the estimates based on the full range. Moreover, the predictive risk was about three times higher in studies that used a latency time of 10 years (RR=1.030 for 4 f-y/ml and RR=1.329 for 40 f-y/ml) compared to studies that used no latency (RR=1.012 for 4 f-y/ml and RR= 1.126 for 40 f-y/ml) between lung cancer and exposure (Appendix Table A2, after correction for intercept).

Fibre type

After stratification of the results by fibre type, we observed a non-significant 3 to 4-fold higher combined RR for studies investigating exposure to mixed and amphibole fibres compared to studies investigating exposure predominantly to chrysotile fibres (table 3, model 2A after correction for intercept). Additional analyses showed that these potency differences decreased to about 2-fold at higher exposures (Appendix Table A3, model 2A). The relative potencies across the exposure range are also shown in Figure 3.

Table 3: Predicted risk at different exposure levels stratified by fibre type based on the natural spline

	Intercept (95% CI)	RR 4 f-y/ml (95% CI)	RR 40 f-y/ml (95% CI)
Chrysotile			
natural spline (model 2A)	1.325 (1.115-1.575)	1.334 (1.124-1.583)	1.411 (1.157-1.719)
corrected for intercept	1.000 (0.841-1.188)	1.006 (0.848-1.194)	1.064 (0.873-1.297)
fitted without intercept (model 2B)	1.000 (1.000-1.000)	1.013 (0.999-1.028)	1.142 (0.991-1.315)
Amphiboles			
natural spline (model 2A)	1.888 (1.047-3.402)	1.929 (1.073-3.468)	2.326 (1.297-4.170)
corrected for intercept	1.000 (0.555-1.802)	1.022 (0.568-1.837)	1.232 (0.687-2.209)
fitted without intercept (model 2B)	1.000 (1.000-1.000)	1.109 (1.084-1.134)	2.637 (2.120-3.280)
Mixed			
natural spline (model 2A)	1.291 (0.872-1.912)	1.314 (0.890-1.940)	1.541 (1.065-2.231)
corrected for intercept	1.000 (0.675-1.481)	1.018 (0.690-1.503)	1.194 (0.825-1.727)
fitted without intercept (model 2B)	1.000 (1.000-1.000)	1.028 (1.019-1.038)	1.322 (1.208-1.446)

RR=relative risk

When spline regressions were fitted without intercept (table 3, model 2B), amphiboles had an 8 to 12-fold increased risk compared to chrysotile which was statistically significant. However, the exposure-response relationship based on the spline without intercept seems to be unrealistic and uncertain at higher cumulative exposures for amphiboles since the risk decreased after exposure of 150 f-y/ml (Figure 3).

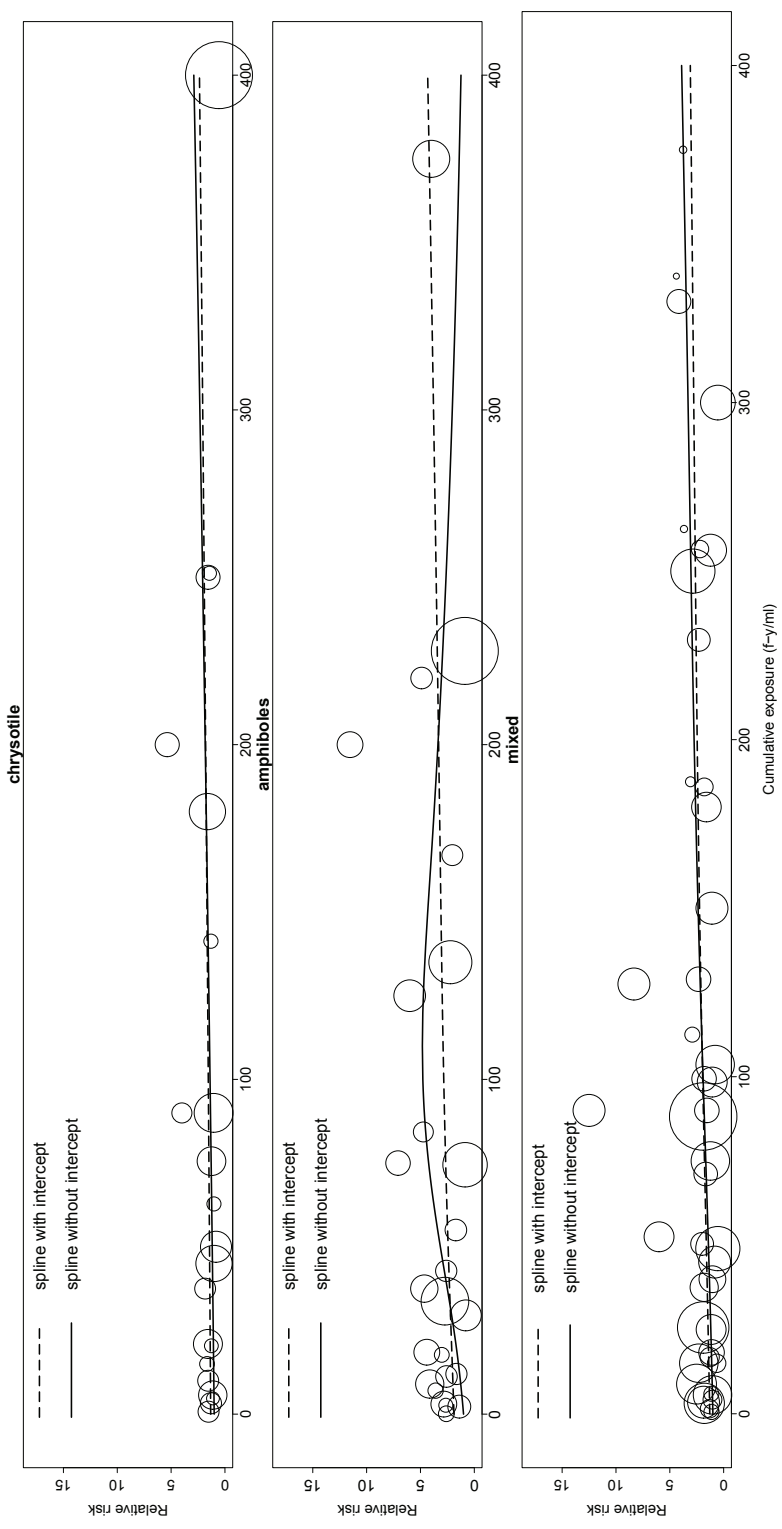


Figure 3 Predicted exposure-response relationship over an exposure range of 0-400 f-y/ml stratified by fibre type (based on a spline regression model fitted with and without intercept)

The predicted risk for chrysotile at low cumulative exposure ranges was heavily influenced by the Quebec and South Carolina studies. For exposures of 4 f-y/ml, the exclusion of Quebec and the South Carolina study yielded a corrected predicted risk of 1.016 and 1.001, respectively, as compared to the overall estimate of 1.006. When both studies were excluded, this risk was estimated to be 1.004. The estimated risks for amphiboles were largely driven by the Wittenoom study (#6) and the New Jersey study (#7). When both studies were removed from the analyses, the risk for exposures of 4 f-y/ml dropped from 1.022 to 1.005. For mixed fibres, the predicted risk was most heavily influenced by the Belgian study (#14). Upon exclusion of the Belgian study, the risk increased from 1.018 to 1.027 for mixed exposures of 4 f-y/ml. Removing these five most influential studies from the analyses resulted in a 1.3 to 7-fold higher combined risk for studies investigating exposure to amphibole and mixed fibres compared to studies investigating exposure predominantly to chrysotile fibres.

Comparison with risk estimates from other meta-analysis

An overview of the risk estimates based on previously published meta-analyses for cumulative exposure estimates of 4 and 40 f-y/ml is shown in Table 4. Our overall point estimates were higher compared to the risks that we calculated based on the meta-K_L value presented for the same 19 studies in the study of Lenters et al..[4] Under a random linear effect model, they observed a K_L value (*100) of 0.13 (95% CI: 0.04-0.22) with an intercept of 1.47 when al 19 studies were considered. Also, they showed that the meta-K_L value was higher for studies with a better exposure measurement strategy. When studies with two or more limitations in the exposure assessment component were excluded, their meta- K_L value was about two times higher. An ad hoc analysis showed that predictions based on our model also yielded higher risk estimates for studies with fewer limitations in the exposure assessment component (Appendix Table A4).

Estimates based on the overall meta-K_L value from the meta-analysis of Lash et al.[5] were similar albeit slightly lower. Overall estimates based on the analyses of Berman and Crump were comparable to our estimates.[2] However, they used a proxy for PCM measurements based on transmission electron microscopy (TEM), which complicates direct comparisons. If analyses would have been performed with results based on PCM measurements, their estimates would have been considerably lower (Supplemental Material, Table 6 of the meta-analysis of Lenters et al.).[4] Although we observed higher or comparable overall risks, we observed a much lower potency difference between amphiboles and chrysotile compared to those observed by Berman and Crump, and Hodgson and Darnton.[2;3]

Table 4 Overview of predicted risk based on previously published meta-analyses

Predictions	K_L*100	RR 4 f-y/ml	RR 40 f-y/ml
Spline (results of this study)			
overall (corrected for intercept)		1.013	1.133
amphiboles (corrected for intercept)		1.022	1.232
chrysotile (corrected for intercept)		1.006	1.064
mixed (corrected for intercept)		1.018	1.194
Based on meta-analysis of Lenters et al. (for fibres measured by PCM)[4] [a]			
overall	0.13	1.0052	1.052
amphiboles	0.33	1.0132	1.132
chrysotile	0.04	1.0016	1.016
mixed	0.13	1.0052	1.052
Based on meta-analysis of Lash et al. (for fibres measured by PCM)[5] [b]			
overall	0.26	1.0104	1.104
Based on meta-analysis of Hodgson and Darnton (for fibres measured by PCM as assigned by TEM)[3] [c]			
amphiboles	5 [d]	1.097	2.936
chrysotile	0.1 [d]	1.002	1.034
mixed	0.32 [d]	N.A. [d]	N.A. [d]
Based on meta-analysis of Berman and Crump[2] [e]			
for fibres measured by PCM			
overall (when restricting the relative potency of chrysotile to amphibole to 1) [f]	0.34	1.0136	1.136
for long fibres (length >10 µm) of al width measured by TEM			
amphiboles	2.7	1.108	2.08
chrysotile	0.29	1.0116	1.116
for long fibres (length >10 µm) with width <4µm measured by TEM			
amphiboles	7.7	1.308	4.08
chrysotile	0.49	1.0196	1.196

RR=relative risk; K_L= the excess relative risk per unit of fibre year; PCM=phase contrast microscopy and measures fibres of longer than 5µm, thicker than approximately 0.25 µm, and with an aspect (length-to-width) ratio >3. TEM=transmission electron microscopy.

[a] Estimates were based on a random effect model by combining K_L values that were derived by fitting an additive linear risk model with a variable intercept to each study;

[b] Overall estimates were based on a random effect model by combining K_L values that were derived by fitting an additive linear risk model to each study (the K_L values and the intercepts were assumed to have a log normal distribution).

[c] Estimates were based on exposure-risk relationships across cohorts by calculating an average exposure and an excess risk for each cohort;

[d] K_L values shown in the table are based on moderate or higher exposures. For low exposures, risks were calculated by applying the sub-linear model: RR=1.6*cumulative exposure^1.3 for amphiboles and RR=0.028*cumulative exposure^1.3 for chrysotile (as assessed by the authors for the best fitted model). No model for low exposures of mixed fibres was assessed.

[e] Estimates were based by fitting K_L-values and matching fibre type and size dimensions (as determined by TEM). The K_L values were derived by fitting an additive linear risk model with a variable intercept (with a maximum of RR=2) to each study;

[f] The K_L value based on PCM was assessed by fitting a metric with fibres of >0.2 µm width in which the relative potencies of long fibres versus short fibres and chrysotile versus amphibole were restricted to 1.

Discussion

We used all available quantitative exposure-response data from observational epidemiological studies to assess the association between cumulative asbestos exposure and the risk of lung cancer. Our estimates for low level exposures are of particular interest to predict the impact of exposures on individuals occupationally exposed to low levels and the general population. We estimated the RR for lung cancer to be 1.013 (95% CI: 0.791-1.296) for 4 f-y/ml and 1.133 (95% CI: 0.888-1.444) for 40 f-y/ml cumulative exposure. These predictions were based on a natural spline model that best fitted our data. When no intercept was fitted, significantly higher RRs were observed ranging from 1.027 (95% CI: 1.020-1.034) for 0.4 f-y/ml to 1.301 (95% CI: 1.215-1.392) for 40 f-y/ml. Our most conservative predicted risks were equal or higher than estimates based on additive linear relative risk models applied in previously published meta-analysis.[4;5] Furthermore, our results indicated a moderately higher increased risk at low exposure in studies investigating amphiboles and mixed fibres compared to studies investigating chrysotile. These potency differences, however, were strongly influenced by a few studies. In general, we observed a lower potency difference between fibre types compared to those observed in previous meta-analyses.[2;3]

The large heterogeneity between individual study results motivated the use of a random intercept and slope model consistent with previous meta-analyses.[2-5] The natural spline model provided the best fit to the data. After retransforming the results to the original scale our results substantiated the evidence that the RR increases virtually linear with increasing exposure. Our findings are in contrast to data of Hodgson and Darnton suggesting a sub-linear relationship.[3] One might also have expected a more supra-linear effect based on substantial high risks observed at very low exposures in a population based study.[7] Although this population based study was included in the current meta-regression our results were statistically compatible with a more-or-less linear exposure-response model. The advantage, however, of our new method is that it provides a more accurate estimate of the lung cancer risk at low exposure since all available information could be used, and estimates did not need to be based on extrapolations below the study-specific exposure range. Moreover, our predictions are not heavily dependent on estimates at high exposure levels which are vulnerable to measurement error.[18] Substantially higher risk at low exposure has been observed in population based studies with semi-quantitative results.[19;20] Although estimates from these studies are quite high, they are in the range of our results when we included only high quality studies.

We adjusted the predicted estimates for the intercept. This assumes that the intercept fully represents a difference in baseline risk and in practice this may not be true. In fact, the observed intercepts suggest a very high excess risk (about 50%) among workers compared to the general population that is attributable to other factors than asbestos exposure. Besides differences in risk factors between the exposed and unexposed population, systematic and

random measurement errors can lead to an intercept greater than one.[13] In the study of Lenters et al., a critical review was performed on the quality of the exposure assessment methodology of the included studies.[4] Here it was shown that only a few studies had few limitations in the exposure assessment component and were of high quality. Furthermore, studies with lower quality had on average higher intercepts. Therefore, it is reasonable to assume that the observed intercept above RR=1 is at least partly due to measurement error. Therefore, we also showed results of the natural spline fitted without intercept, since one might suggest that fitting a line through the origin (ln(RR)=0 at zero exposure) would be more appropriate in the case of random measurement error. Our results also showed a lower intercept for the studies that included a latency time compared to those that did not include a latency time. Including no latency between exposure and lung cancer could also lead to measurement error in the exposure assessment when it incorrectly reflects the etiological time window of exposure.

Low cumulative exposures are associated with all kind of occupations if duration of exposure is short. However, low cumulative exposures have been particularly observed in the general population due to downstream use of asbestos.[7] Like other meta-analyses, we could not determine whether risks might differ by exposure intensities, since intensities could mostly not be distinguished from reported cumulative exposures. Information on intensities is especially important if a threshold exists for asbestos related lung cancer. However, no threshold of exposure intensity has been delineated for asbestos related lung cancer. Moreover, a study by Frost et al.[21] showed that long term asbestos removal workers had a significant increased risk of lung cancer compared to short term workers indicating that cumulative exposure is an important measure if persons are exposed to low intensities.

The degree to which different types of asbestos have different potencies is a topic of ongoing debate.[2;22] Berman and Crump showed a 9 times higher increased risk for long amphiboles compared to long chrysotile fibres of all widths, and had even higher estimates for specific diameters (a ratio of 16:1 for long amosite versus long chrysotile for fibres with widths <4 µm).[2;11] Hodgson and Darnton estimated the risk difference between chrysotile and amphibole fibres for lung cancer to be between 10 and 50.[3] In the study of Lenters et al., a difference in risk ratio of a factor 8 was observed when all 19 studies were included (i.e. without adjusting for quality).[4] In our analyses, we observed a non-significant 3 to 4-fold difference in potency between chrysotile and amphibole fibres.

Various explanations exist for the higher potency differences observed in previous mete-analyses compared to our results. Firstly, we used non-linear regressions, and estimated the overall slope from a distribution of study slopes. This resulted in shrinkage of study specific slopes to the overall combined slope as well as less weight of point estimates at high exposures, and therefore, our analyses are less influenced by extreme results. Secondly,

among the amphibole studies, we observed very high intercepts for the Wittenoom and New Jersey studies (i.e. intercepts of 2.8 and 3.8, respectively). These high intercepts were partly due to very high risks observed at relatively low exposures: the Wittenoom study observed a risk of 2.6 for 0.11 f-y/ml and the New Jersey study a risk of 2.8 for 3 f-y/ml. In the meta-analysis by Berman and Crump these high intercepts were truncated at 2.[2;12] Therefore, our estimated risk for amphiboles is likely to be lower compared to risks estimated by Berman and Crump. When we fitted a natural spline without intercept, we observed a significant increased risk for amphiboles. In this case, the ratio of potency for amphobiles versus chrysotile was estimated to lie between 8:1 and 12:1, which was comparable to the ratios observed in the analyses of Berman and Crump for long fibres. However, the observed exposure-response relation for amphiboles based on the spline without intercept was uncertain at higher cumulative exposure levels. Thirdly, Berman and Crump controlled for different fibre sizes in their meta-analysis.[2] Several studies showed that relatively longer and thinner fibres are stronger associated with lung cancer.[2;23;24] Since chrysotile fibres are generally longer and thinner than amphiboles, this might also explain the higher potency ratio between fibres types observed by Berman and Crump. Finally, Hodgson and Darnton used a different methodology to estimate the asbestos-related lung cancer risk.[3] They derived exposure-risk relationships across cohorts by calculating an average exposure and an excess risk for each cohort to avoid the effect of random measurement error. However, when for example misclassification is more severe in lower exposure categories, the method applied does not necessarily completely eliminate the effect of exposure misclassification. Furthermore, mean levels do not reflect actual exposure levels accurately when observations are skewed. Also, it is expected that extraneous risk factors are differential distributed across study cohorts, which can have influenced their results.

From our results it was apparent that the Quebec mine study and South Caroline textile study had a significant impact on the risk estimates for chrysotile. Upon removing the Quebec study, the RR for chrysotile increased considerably, whereas the exclusion of the South Carolina resulted in lower risks for chrysotile. The combined estimate of the three other studies involving chrysotile exposure also showed relatively low risks. The differences between the Quebec mine and South Carolina textile studies have been discussed extensively. A recent study by Berman concluded that the characteristics of the fibre can potentially explain the differences in lung cancer potency observed between these cohorts.[25] In that study, it was shown that the South Carolina textile workers were exposed to longer asbestos structures compared to the Quebec miners and millers. The PCM-counted structures in textile factory dusts were virtually 100% asbestos and 100% asbestiform. In contrast, at least one third of the structures counted by PCM in chrysotile mine and mill dusts were not asbestos. Additional limitations of PCM measurements have been discussed elsewhere.[26] Interestingly, the South Carolina study was classified as one with no limitations and the Quebec study as one with several limitations in the exposure assessment component as

assessed by Lenters et al.[4] They showed that better quality studies yielded higher meta-estimates. This pattern was also observed with our spline regression model suggesting that observed fibre specific potency differences at low cumulative exposure might also be partly due to differences in quality. Moreover, Lenters at al. showed that when analysis is restricted to only studies with few quality limitations of the exposure assessment component, the epidemiological evidence base is too sparse to draw deductions about potency differences per fibre type. Therefore, in light of the quality, we could not easily ascertain the magnitude of the potency differences between different fibres at low cumulative exposure.

Conclusion

Our results showed relative lung cancer risks for asbestos exposures of 4 f-y/ml, and 40 f-y/ml to be between 1.013 and 1.027, and 1.13 and 1.30, respectively. Although we could not unequivocally determine potency differences between different fibre types at very low exposure levels of asbestos, the collected evidence suggests a 3-fold difference in risk between chrysotile and amphibole asbestos. This potency difference was not significant and lower than the generally held consensus. The flexible spline regression model we applied indicated that for low cumulative exposures, the increase in relative risk of lung cancer due to asbestos exposure may be larger than expected from previous results. This would suggest that, in general, a larger fraction of lung cancer incidence may be attributable to (many individuals having) relatively low cumulative exposure levels than previously estimated and might have important implications in developed nations. Additional research is required, in particular among removal workers and the general population in developed countries or low exposed industrial cohorts using quantitative estimates of asbestos exposure, to further substantiate this notion.

Acknowledgements

This study was supported by the Institute for Asbestos Victims, the Netherlands.

Conflict of interest

The authors declare that they have no conflict of interest.

Reference List

1. Straif K, brahim-Tallaa L, Baan R, Grosse Y, Secretan B, El GF et al. A review of human carcinogens--part C: metals, arsenic, dusts, and fibres. Lancet Oncol 2009; 10:453-454.
2. Berman DW, Crump KS. A meta-analysis of asbestos-related cancer risk that addresses fiber size and mineral type. Crit Rev Toxicol 2008; 38 Suppl 1:49-73.
3. Hodgson JT, Darnton A. The quantitative risks of mesothelioma and lung cancer in relation to asbestos exposure. Ann Occup Hyg 2000; 44:565-601.
4. Lenters V, Vermeulen R, Dogger S, Stayner L, Portengen L, Burdorf A et al. A Meta-Analysis of Asbestos and Lung Cancer: Is Better Quality Exposure Assessment Associated with Steeper Slopes of the Exposure-Response Relationships? Environ Health Perspect 2011.
5. Lash TL, Crouch EA, Green LC. A meta-analysis of the relation between cumulative exposure to asbestos and relative risk of lung cancer. Occup Environ Med 1997; 54:254-263.
6. Virta RL. Mineral commodity profiles-Asbestos: U.S. Geological Survey Circular 1255-KK. 56. 2005.
7. Gustavsson P, Nyberg F, Pershagen G, Scheele P, Jakobsson R, Plato N. Low-dose exposure to asbestos and lung cancer: dose-response relations and interaction with smoking in a population-based case-referent study in Stockholm, Sweden. Am J Epidemiol 2002; 155:1016-1022.
8. Vlaanderen J, Portengen L, Rothman N, Lan Q, Kromhout H, Vermeulen R. Flexible meta-regression to assess the shape of the benzene-leukemia exposure-response curve. Environ Health Perspect 2010; 118:526-532.
9. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7:177-188.
10. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992; 135:1301-1309.
11. Harrel FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001.
12. Berman DW, Crump KS. Update of potency factors for asbestos-related lung cancer and mesothelioma. Crit Rev Toxicol 2008; 38 Suppl 1:1-47.
13. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med 1998; 55:651-656.
14. Shi JQ, Copas JB. Meta-analysis for trend estimation. Stat Med 2004; 23:3-19.
15. National Institute for Occupational Safety and Health. *Revised Recommended Asbestos Standard.* Washington, DC: National Institute for Occupational Safety and Health. 1976. (DHHS (NIOSH) Publication No. 77-169).
16. Occupational Safety and Health Administration Health. *Asbestos Standard For General Industry.* Washington, DC: Occupational Safety and Health Administration. 1995. (OSHA publication 3095 1995 (revised)).
17. Rothpearl A. The jackknife technique in statistical analysis. Chest 1989; 95:940.
18. Stayner L, Steenland K, Dosemeci M, Hertz-Picciotto I. Attenuation of exposure-response curves in occupational cohort studies at high exposure levels. Scand J Work Environ Health 2003; 29:317-324.
19. van Loon AJ, Kant IJ, Swaen GM, Goldbohm RA, Kremer AM, van den Brandt PA. Occupational exposure to carcinogens and risk of lung cancer: results from The Netherlands cohort study. Occup Environ Med 1997; 54:817-824.
20. De Matteis S., Consonni D, Lubin JH, Tucker M, Peters S, Vermeulen RC et al. Impact of occupational carcinogens on lung cancer risk in a general population. Int J Epidemiol 2012; 41:711-721.
21. Frost G, Harding AH, Darnton A, McElvenny D, Morgan D. Occupational exposure to asbestos and mortality among asbestos removal workers: a Poisson regression analysis. Br J Cancer 2008; 99:822-829.
22. Stayner LT, Dankovic DA, Lemen RA. Occupational exposure to chrysotile asbestos and cancer risk: a review of the amphibole hypothesis. Am J Public Health 1996; 86:179-186.
23. Loomis D, Dement J, Richardson D, Wolf S. Asbestos fibre dimensions and lung cancer mortality among workers exposed to chrysotile. Occup Environ Med 2010; 67:580-584.
24. Stayner L, Kuempel E, Gilbert S, Hein M, Dement J. An epidemiological study of the role of chrysotile asbestos fibre dimensions in determining respiratory disease risk in exposed workers. Occup Environ Med 2008; 65:613-619.
25. Berman DW. Comparing milled fiber, Quebec ore, and textile factory dust: has another piece of the asbestos puzzle fallen into place? Crit Rev Toxicol 2010; 40:151-188.
26. Dement JM, Kuempel ED, Zumwalde RD, Smith RJ, Stayner LT, Loomis D. Development of a fibre size-specific job-exposure matrix for airborne asbestos fibres. Occup Environ Med 2008; 65:605-612.

27. Liddell FD, McDonald AD, McDonald JC. The 1891-1920 birth cohort of Quebec chrysotile miners and millers: development from 1904 and mortality to 1992. Ann Occup Hyg 1997; 41:13-36.
28. Pira E, Pelucchi C, Piolatto PG, Negri E, Bilei T, La VC. Mortality from cancer and other causes in the Balangero cohort of chrysotile asbestos miners. Occup Environ Med 2009; 66:805-809.
29. McDonald AD, Fry JS, Woolley AJ, McDonald JC. Dust exposure and mortality in an American chrysotile asbestos friction products plant. Br J Ind Med 1984; 41:151-157.
30. Hein MJ, Stayner LT, Lehman E, Dement JM. Follow-up study of chrysotile textile workers: cohort mortality and exposure-response. Occup Environ Med 2007; 64:616-625.
31. Loomis D, Dement JM, Wolf SH, Richardson DB. Lung cancer mortality and fibre exposures among North Carolina asbestos textile workers. Occup Environ Med 2009; 66:535-542.
32. Berry G, de Klerk NH, Reid A, Ambrosini GL, Fritschi L, Olsen NJ et al. Malignant pleural and peritoneal mesotheliomas in former miners and millers of crocidolite at Wittenoom, Western Australia. Occup Environ Med 2004; 61:e14.
33. Seidman H, Selikoff IJ, Gelb SK. Mortality experience of amosite asbestos factory workers: dose-response relationships 5 to 40 years after onset of short-term work exposure. Am J Ind Med 1986; 10:479-514.
34. Levin JL, McLarty JW, Hurst GA, Smith AN, Frank AL. Tyler asbestos workers: mortality experience in a cohort exposed to amosite. Occup Environ Med 1998; 55:155-160.
35. Sullivan PA. Vermiculite, respiratory disease, and asbestos exposure in Libby, Montana: update of a cohort mortality study. Environ Health Perspect 2007; 115:579-585.
36. Berry G, Newhouse ML. Mortality of workers manufacturing friction materials using asbestos. Br J Ind Med 1983; 40:1-7.
37. Finkelstein MM. Mortality among employees of an Ontario asbestos-cement factory. Am Rev Respir Dis 1984; 129:754-761.
38. Hughes JM, Weill H, Hammad YY. Mortality of workers employed in two asbestos cement manufacturing plants. Br J Ind Med 1987; 44:161-174.
39. Albin M, Jakobsson K, Attewell R, Johansson L, Welinder H. Mortality and cancer morbidity in cohorts of asbestos cement workers and referents. Br J Ind Med 1990; 47:602-610.
40. Lacquet LM, van der Linden L, Lepoutre J. Roentgenographic lung changes, asbestosis and mortality in a Belgian asbestos-cement factory. IARC Sci Publ 1980;783-793.
41. Enterline PE, Hartley J, Henderson V. Asbestos and cancer: a cohort followed up to death. Br J Ind Med 1987; 44:396-401.
42. Selikoff IJ, Seidman H. Asbestos-associated deaths among insulation workers in the United States and Canada, 1967-1987. Ann N Y Acad Sci 1991; 643:1-14.
43. McDonald AD, Fry JS, Woolley AJ, McDonald JC. Dust exposure and mortality in an American factory using chrysotile, amosite, and crocidolite in mainly textile manufacture. Br J Ind Med 1983; 40:368-374.
44. Peto J, Doll R, Hermon C, Binns W, Clayton R, Goffe T. Relationship of mortality to measures of environmental asbestos pollution in an asbestos textile factory. Ann Occup Hyg 1985; 29:305-355.

Appendix

Table A1: Predicted risk at different exposure levels: based on sensitivity analyses excluding risk estimates corresponding to exposure categories of > 100 f-y/ml*

	Deviance (df) [a]	Intercept [b] (95% CI)	RR 4 f-y/ml [b] (95% CI)	RR 40 f-y/ml [b] (95% CI)
Models*				
1A. linear model	49.9 (56)	1.429 (1.133-1.803)	1.454 (1.156-1.830)	1.703 (1.327-2.184)
corrected for intercept		1.000 (0.793-1.262)	1.018 (0.809-1.280)	1.192 (0.929-1.529)
1B. linear model without intercept	399.5 (58)	1.000 (1.000-1.000)	1.050 (1.023-1.077)	1.626 (1.259-2.099)
2A. natural spline	49.7 (55)	1.454 (1.136-1.862)	1.472 (1.162-1.864)	1.676 (1.292-2.175)
corrected for intercept		1.000 (0.781-1.280)	1.012 (0.799-1.282)	1.152 (0.888-1.495)
2B. natural spline without intercept	330.1 (57)	1.000 (1.000-1.000)	1.117 (1.085-1.149)	2.173 (1.690-2.795)

RR=relative risk; df=degrees of freedom calculated as the number of data points minus the number of coefficients estimated; [a] fitted using ML estimation; [b] fitted using REML estimation; *The deviance of the empty and intercept only model was 3433.9 and 309.9 respectively. *Risk predictions were based on models that were fitted to risk estimates corresponding to exposures of 100 f-y/ml or less.

Table A3: Predicted risk at high exposure levels stratified by fibre type based on the natural spline

	RR 250 f-y/ml (95% CI)	RR 400 f-y/ml (95% CI)
Chrysotile		
natural spline (model 2A)	1.925 (0.908-4.082)	2.367 (0.720-7.784)
corrected for intercept	1.452 (0.685-3.080)	1.786 (0.543-5.873)
fitted without intercept (model 2B)	2.074 (0.860-5.003)	2.891 (0.708-11.81)
Amphiboles		
natural spline (model 2A)	3.541 (1.816-6.905)	4.330 (1.884-9.949)
corrected for intercept	1.876 (0.962-3.658)	2.294 (0.998-5.270)
fitted without intercept (model 2B)	2.627 (0.770-8.965)	1.237 (0.171-8.961)
Mixed		
natural spline (model 2A)	2.627 (1.713-4.027)	3.097 (1.860-5.155)
corrected for intercept	2.034 (1.327-3.119)	2.398 (1.441-3.992)
fitted without intercept (model 2B)	3.038 (1.797-5.138)	3.905 (1.684-9.053)

RR=relative risk

Table A2: Investigating the influence of inclusion of a latency between exposure and lung cancer on the predicted risks based on the natural spline

Exclusion	Intercept	RR 4 f-y/ml	RR 40 f-y/ml	Studies included
All 19 studies	1.483 (1.157-1.900)	1.502 (1.173-1.922)	1.680 (1.317-2.142)	1-19
corrected for intercept	1.000 (0.780-1.281)	1.013 (0.791-1.296)	1.133 (0.888-1.444)	
fitted without intercept	1.000 (1.000-1.000)	1.027 (1.020-1.034)	1.301 (1.215-1.392)	
No latency (i.e. unlagged exposure)	1.672 (1.111-2.515)	1.692 (1.127-2.539)	1.882 (1.270-2.789)	2,3,6,7,8,10,11,13, 14,15,19
corrected for intercept	1.000 (0.664-1.504)	1.012 (0.674-1.519)	1.126 (0.760-1.668)	
fitted without intercept	1.000 (1.000-1.000)	1.037 (1.026-1.049)	1.436 (1.286-1.605)	
Latency time of 10 years (i.e. exposures in the 10 years previous to follow-up were discarded)	1.144 (0.855-1.532)	1.178 (0.884-1.569)	1.520 (1.159-1.994)	4,5,9,12,16,17
corrected for intercept	1.000 (0.747-1.339)	1.030 (0.773-1.372)	1.329 (1.013-1.743)	
fitted without intercept	1.000 (1.000-1.000)	1.041 (1.031-1.051)	1.483 (1.351-1.628)	
Latency time other [a]	1.169 (1.042-1.312)	1.176 (1.050-1.317)	1.236 (1.081-1.413)	1,18
corrected for intercept	1.000 (0.891-1.122)	1.006 (0.898-1.127)	1.057 (0.925-1.209)	
fitted without intercept	1.000 (1.000-1.000)	1.009 (0.998-1.019)	1.090 (0.980-1.212)	

RR=relative risk; [a] In study 1 cumulative exposure was assessed up to age 55 and in study 18 a latency time of 5 years was used (i.e. exposures in the 5 years previous to follow-up were discarded).

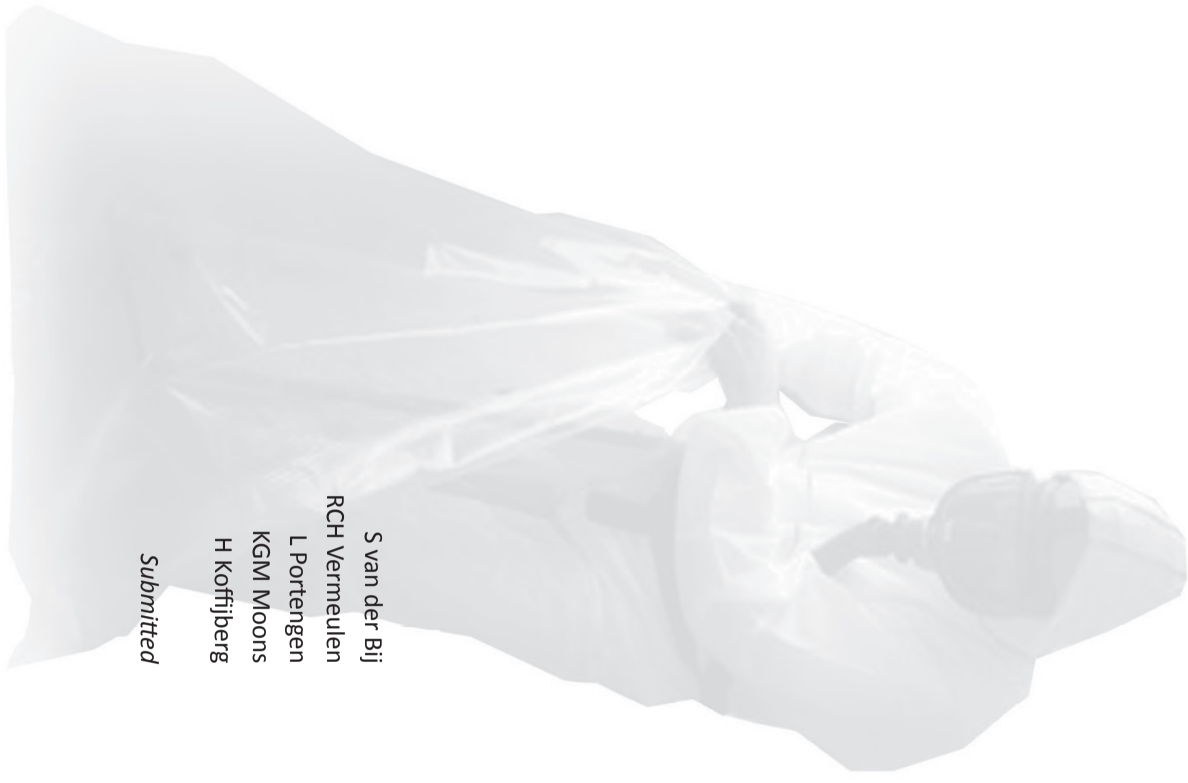
Table A4: Investigating the influence of the quality of the exposure assessment on the predicted risks*

Exclusion	Intercept	RR 4 f-y/ml	RR 40 f-y/ml	Studies included
All 19 studies	1.483 (1.157-1.900)	1.502 (1.173-1.922)	1.680 (1.317-2.142)	1-19
corrected for intercept	1.000 (0.780-1.281)	1.013 (0.791-1.296)	1.133 (0.888-1.444)	
fitted without intercept	1.000 (1.000-1.000)	1.027 (1.020-1.034)	1.301 (1.215-1.392)	
studies with 3-5 limitations scored (lowest quality)	1.818 (1.215-2.719)	1.834 (1.227-2.742)	1.986 (1.331-2.962)	1,2,3,6,7,11,13,15,16,17
corrected for intercept	1.000 (0.668-1.496)	1.009 (0.675-1.508)	1.092 (0.732-1.629)	
fitted without intercept	1.000 (1.000-1.000)	1.024 (1.013-1.035)	1.270 (1.143-1.410)	
studies with 0-2 limitations scored	1.243 (0.985-1.568)	1.261 (1.004-1.585)	1.437 (1.149-1.796)	4,5,8,9,10,12,14,18,19
corrected for intercept	1.000 (0.792-1.261)	1.014 (0.808-1.275)	1.156 (0.924-1.445)	
fitted without intercept	1.000 (1.000-1.000)	1.030 (1.018-1.041)	1.330 (1.193-1.483)	
studies with 0-1 limitations scored	1.288 (1.061-1.564)	1.320 (1.096-1.589)	1.630 (1.333-1.994)	4,5,9,18
corrected for intercept	1.000 (0.824-1.214)	1.025 (0.851-1.234)	1.266 (1.035-1.548)	
fitted without intercept	1.000 (1.000-1.000)	1.047 (1.030-1.065)	1.552 (1.321-1.824)	
no scored limitations (highest quality)	1.377 (1.153-1.644)	1.437 (1.226-1.684)	2.012 (1.537-2.633)	4,9
corrected for intercept	1.000 (0.837-1.194)	1.044 (0.890-1.223)	1.461 (1.116-1.912)	
fitted without intercept	1.000 (1.000-1.000)	1.085 (1.053-1.118)	2.019 (1.541-2.645)	

RR=relative risk; *Results are based on the natural spline model fitted with intercept. The quality of the exposure assessment used in the studies was evaluated in the meta-analysis of Lenters et al. and limitations were scored based on five criteria: 1) sufficient documentation, 2) ratio of highest: lowest cumulative exposure category >50, 3) internal conversion factor (mppcf-y to f-y/ml), 4) coverage of exposure data>30% of exposure history, 5) accurate job histories (for more details see study of Lenters et al. (2001))

Chapter 6

The burden of asbestos-related lung cancer: a comparison of methods



S van der Bij
RCH Vermeulen
L Portengen
KGM Moons
H Koffijberg

Submitted

Abstract

Introduction: Exposure to asbestos fibres is known to increase the risk of mesothelioma and lung cancer. While the vast majority of mesothelioma cases are generally accepted as being caused by asbestos, the proportion of asbestos-related lung cancers is less clear and cannot be determined directly because cases are not clinically distinguishable from those due to other causes. Various modelling methods may be applied to estimate the expected future number of lung cancers due to past and current asbestos exposure. We applied three different methods to the Dutch population, and discuss their evidence requirements, (dis) advantages, and the (dis)similarity of their results.

Methods: We compared three methods that differ in complexity and required evidence. The first method was relatively simple and required little evidence, estimating asbestos-related lung cancer cases directly from observed and predicted mesothelioma cases in an Age-Period-Cohort (APC) analysis. The second method required evidence on the fraction of lung cancer cases attributable (PAR) to asbestos exposure. The third method was the most comprehensive, requiring actual exposure information to perform a life table analysis. Input parameters in the life table analysis were first calibrated using observed mesothelioma cases.

Results: In our analysis of the number of future asbestos-related lung cancer cases in The Netherlands we found that the three methods produced very different estimates: APC method 17,500-22,150 cases, PAR method 12,150 cases, the life table method (life table analyses) 6,300 cases.

Conclusion: The preferred method for estimating asbestos-related lung cancer cases depends on the evidence that is available and the accuracy of this information. We show that using three different methods results in different absolute estimates varying by a factor of ~3.5. As such the exact impact of asbestos exposure on the lung cancer burden remains uncertain.

Introduction

Exposure to asbestos is known to increase the risk of developing mesothelioma and lung cancer.[1] While the vast majority of mesothelioma cases are generally accepted as being caused by asbestos, the proportion of asbestos-related lung cancers is less clear and cannot be determined directly because cases are not clinically distinguishable from those due to other causes.[2] Consequently, the historical number of asbestos-related lung cancers is unknown and cannot be used to forecast future number of asbestos-related lung cancers. This means that any prediction of the future number of asbestos-related lung cancers in the general population is necessarily based on mathematical models.

To estimate the future number of asbestos-related lung cancers in the general population, the specific model structure depends on the available evidence and preference of the researchers. Life tables analyses can be used to estimate the future number of asbestos-related cancers.[3;4] However, due to limited availability of estimates of asbestos exposure in the general population these analyses are often limited to cohort studies of asbestos exposed workers.[5] Hence, alternative models might need to be selected to estimate the number of asbestos-related lung cancers in the general population. An alternative approach is a model based on the population attributable risk (PAR) which can be derived from lung cancer case control or cohort studies[6-8]. Another method is the use of forecasted mesothelioma cases and to translate these estimates in a prediction of future lung cancer cases based on observed ratio's of these two cancers in asbestos exposed populations.[2;9-12]

Although any model that is applied may provide estimates of uncertainty, this uncertainty indicates only how the result of the analysis may vary given the uncertainty in the input values for the model, assuming the model structure itself is correctly specified. This assumption, however, cannot usually be verified. When the correct model structure is unknown, structural uncertainty can and should be assessed by comparing the results from different model structures.

In this paper we compare three model structures for the prediction of the expected future asbestos-related lung cancers in the Netherlands in the period 2011-2030. The first model estimated the asbestos-related lung cancer cases directly from predicted mesothelioma cases, the second model made use of the PAR and the third model used life table analyses. We discuss their advantages, disadvantages and evidence requirements, and compare their results.

Methods

Three methods were applied to predict the number of future asbestos-related lung cancers in the Netherlands in the period 2011-2030. All methods were applied separately to men and women and results were aggregated.

Model 1: the APC model

Estimates of asbestos-related lung cancer cases can be derived directly from predicted mesothelioma cases through a conversion factor.[12] To predict the number of mesothelioma cases an Age-Period-Cohort (APC) model was constructed.[13;14] Required data were provided by Statistics Netherlands (CBS) and included the observed number of mesothelioma deaths in 1969-2010, annual demographic distributions of the Dutch population in 1969-2010, and expected demographic distributions for 2011-2030.[15] Deaths due to mesothelioma were identified by ICD-8/9 code 163.0 (pleural cancer) for the years 1969-1995. As these categories did not include non-pleural mesotheliomas we added an extra 5% to the numbers as the number of non-pleural mesothelioma deaths is estimated to be around 5% of all mesothelioma cases [16-18]. For 1996-2010 the number of mesothelioma was identified by ICD-10 code C45. All data were tabulated into 13 age groups (31-35,36-40,…,86-90, and 91-95) and 8 five year periods following the years 1969-1970 (1971-1975,1976-1980,…,2006-2010). This resulted in 20 partially overlapping 10 year birth cohorts (1876-1885,…,1966-1975, 1971-1980) and one of six years (1874-1880) identified by midpoint year (thus the birth cohort of 1965 comprised those born between 1961-1970). Using the number of mesothelioma deaths observed in 1969-2010 we then calculated age specific mortality rates and cohort relative risks by year of birth using an APC model, separately for men and women. Since pleural mesothelioma under the age of 40 years was very rare the birth cohort of 1965 was the youngest cohort for which a reliable risk estimate could be obtained. As asbestos use after 1984 was very limited and an asbestos ban was implanted in 1993 in the Netherlands,[19;20] birth cohorts beyond 1965 were assigned zero risk of mesothelioma and lung cancer due to asbestos exposure. Estimated age specific rates of mesothelioma per birth cohort were projected on the expected future demographic distributions to predict the future number of mesothelioma deaths. A sensitivity analysis was performed in which the birth cohort of 1970 was assigned the risk of the birth cohort of 1965 instead of zero risk and birth cohorts beyond 1970 were assigned zero risk, to simulate longer propagation of risk over time.

To estimate the future number of lung cancers we used the ratios between mesothelioma and lung cancer from a published meta-analysis.[12] In the Netherlands different type of asbestos have been used, therefore we applied ratios reported for mixed asbestos fibres: an unadjusted ratio of 1.9 (95%CI: 1:1.4-2.6) and a smoking adjusted ratio of 1.5 (95%CI: 1.1-2.0) asbestos-related lung cancers per mesothelioma death. Here, the smoking adjusted ratio is likely to be an underestimation.[12]

Model 2: the PAR model

This model uses the number of lung cancers observed in 2010, and the estimated population attributable risk (PAR) by age categories. In a Dutch study it was estimated that 11.6% of the lung cancers cases that occurred in men of 55-73 years of age in the period 1986-1990 were related to asbestos exposure.[8] This PAR was assumed also to be applicable to the total

number lung cancers in men of > 40 years of age in 2010 and was assumed to be fixed in future years. Persons of younger ages (i.e born after 1970) were assumed never exposed to asbestos and thus the PAR was set to zero. The distribution of lung cancer cases over age in men was derived as the average of the observed distribution over age in the years 2008-2010. When the age categories under consideration contained both persons born before and after 1970 a linear interpolation of the PAR was used. For men, the estimated PAR values over time were then applied to the expected future number of lung cancer cases which was calculated from the observed lung cancer incidence in 2008-2010 and the expected demographic distribution in 2011-2030.[15]

As no reliable PAR estimates are available for Dutch women we first estimated the ratio of the expected number of asbestos-related lung cancers based on the PAR model to the observed number of mesotheliomas among men in 2010. We then applied this ratio to the observed number of mesotheliomas among women in 2010 to derive the number of asbestos-related lung cancers in women in 2010. This resulted in a PAR of 2.5% which is rather similar to the PAR of 2.2% found in a French study applying the same procedure.[6]

Model 3: the life table model

In this model the future number of asbestos-related lung cancers was estimated based on exposure information and the asbestos-related lung cancer risk as a direct function of exposure. To estimate the number of individuals exposed to asbestos we used data from the Netherlands cohort study (NLCS).[21] The NLCS is a prospective cohort study, which started in 1986 among men and women aged 55-69 years (n=120,852). At baseline a comprehensive lifetime job history till 1986 was collected among all participants which was only entered in the computer for selected cases and a randomly drawn subcohort (n=5,000). In total, complete job histories were available of 4,568 participants. We estimated the proportion of participants that were occupationally exposed to asbestos using a general population job-exposure matrix (DOM-JEM). DOM-JEM for asbestos exposure is a semi-quantitative exposure matrix and classifies occupations into *no exposure*, *low exposure* and *high exposure* based on five-digit ISCO-68 codes.[22] All occupations performed before 1945 were assigned to *no exposure*. The DOM-JEM assignment categories were calibrated to an exposure intensity (f/ml) by year (>1945) based on linkage to the EXPOSYN database.[23] The cumulative exposure in fibre years (f-y/ml) subsequently was calculated for each participant by multiplying the exposure intensity by the duration for each recorded job period and then aggregating the exposure estimates over all job periods. As significant occupational asbestos exposure did not occur anymore after 1990, we determined the number of participants exposed up to 1990. For individuals aged 60-74 years in 1990 we used the estimates obtained from participants aged 56-69 years in the NLCS cohort as we assumed that most of them did not have significant asbestos exposure anymore after 1986 due to retirement. For younger age categories estimates of asbestos exposure were determined by using the age specific job distribution of the participants in the NLCS cohort. Finally, the estimated

proportions of exposed participants, estimated from the NLCS cohort by age and gender, were multiplied by the corresponding age and gender specific Dutch population in 1990. Results were extrapolated to the year 2010 and cumulative exposures were then averaged by age and gender.

To estimate the asbestos-related lung cancer risk we estimated age specific lung cancer rates for men and women in the general population. These rates were assessed by Poisson regression using the number of observed lung cancer deaths in 2001-2010.[15] The asbestos-related lung cancer risk was determined by multiplying rates with the relative risk associated with asbestos exposure. This relative risk was determined as RR=1+ K_L*cumulative exposure with K_L = 0.028. The K_L value of 0.028 was observed in cohort of individuals in whom asbestos exposure was quantified by SYN-JEM.[22] To estimate the number of asbestos-related lung cancers standard life table analyses of lung cancer were conducted.[24] Input parameters in the life table analysis were subsequently calibrated using predicted mesothelioma cases. For this, mesothelioma incidence was calculated similar to the US Environmental Protection Agency model with a potency for cumulative exposure (K_m) of 2.53e-8 as estimated for mixed fibre types.[25] We used average estimates for age at first exposure, duration of exposure and cumulative exposure as determined from the NLCS cohort per age category and gender. Subsequently, our estimates of exposure were recalibrated to reproduce the number of mesothelioma cases in 2011 as was observed on average in 2006-2010. Finally, we repeated the full life table analysis with the calibrated exposure data to reassess the expected future number of asbestos-related lung cancers.

Results

The specifications of the three models, in terms of requirements, complexity, underlying assumptions, advantages and disadvantages are summarized in Table 1. This table allows an informal comparison of the models and may be used to check the requirements, and thereby feasibility, of each of the models for application in other settings and countries.

Model 1: APC model

Figure 1 shows the number of future asbestos-related lung cancers as estimated by the APC model, i.e., based on the estimated future number of mesotheliomas between 2011 and 2030. The left panel shows the expected absolute number of asbestos-related lung cancers per year. The right panel shows the corresponding cumulative number of cases per year. The number of asbestos-related lung cancers depended heavily on the applied ratio between mesothelioma and lung cancer. Based on the unadjusted ratio (1:1.9) the number of lung cancers between 2011-2030 was estimated to be around 22,150 whereas the adjusted ratio (1:1.5) resulted in an estimate around 17,500. The sensitivity analyses in which the birth cohort of 1970 was assigned the risk of the birth cohort of 1965 yielded similar results (data not shown). Figure 1 indicates that the annual number asbestos-related-lung cancers is expected to increase up to year 2022 and to decrease thereafter.

Table 1: Characteristics of the applied models

Model	1. Mesothelioma (APC)	2. PAR	3. Life table
Complexity	Low	Low	Moderate
Evidence required	Minimal: -Demographics -Mesothelioma cases -Ratio mesothelioma to asbestos-related lung cancer	Minimal: -Demographics -PAR -Number of lung cancers (single point in time)	Comprehensive: -Demographics -Detailed information about asbestos exposure levels -Incidence of non asbestos-related lung cancer -(Relative) risk of lung cancer from asbestos exposure
Underlying assumptions	-Mesothelioma is proxy for asbestos exposure -Single constant ratio can describe the relation between mesothelioma and lung cancer -Risk of adjacent birth cohorts can be well determined	-The PAR (from a single well designed study) is representative for the total population -The PAR is representative for future years	-Asbestos exposure is representative for the total population -Exposure-response relationship is known -Lung cancer risk does not change over time
Advantages	-Simple to construct -Evidence commonly available	-Simple to construct - Evidence commonly available from case control or case cohort studies	-Gives detailed outcomes and allows for estimation of other statistics (e.g. life expectancies)
Disadvantages	-Ratio between mesothelioma and lung cancer depends heavily on fibre type which decreases robustness of results -It is unknown if ratio may change by changes in asbestos exposure level -It is likely that ratio changes in forecasting due to differences in the dynamics of the disease.	-Does not take into account competing risks -Is a single indicator and does not take into account changes in asbestos and non-asbestos-related lung cancer risk	-Requires evidence that may not be (readably) available - Can easily result in input that may be uncertain as assumptions about the input have to be made

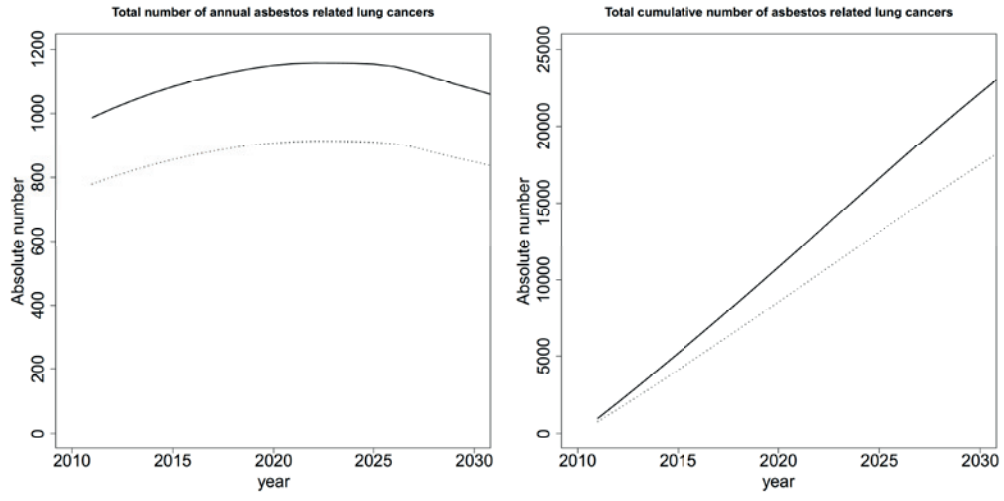


Figure 1: Total number of asbestos-related lung cancer cases in 2011-2030 as estimated by the APC model; i.e., based on the estimated number of mesothelioma between 2011-2030. The solid line shows the results when a ratio between mesothelioma and asbestos-related lung cancer of 1:1.9 was applied, the dashed line shows the results when the smoking adjusted ratio of 1:1.5 was applied.

Model 2: the PAR model

Figure 2 has a layout similar to figure 1, and shows the number of future asbestos-related lung cancers as estimated by the PAR Model, i.e., based on a fixed PAR for the fraction of lung cancers due to asbestos exposure. In this figure, the annual number of asbestos-related lung cancers decreased consistently over time, from 826 in 2011 to 371 in 2030. The cumulative number of asbestos-related lung cancers between 2011 and 2030 was estimated to be around 12,150.

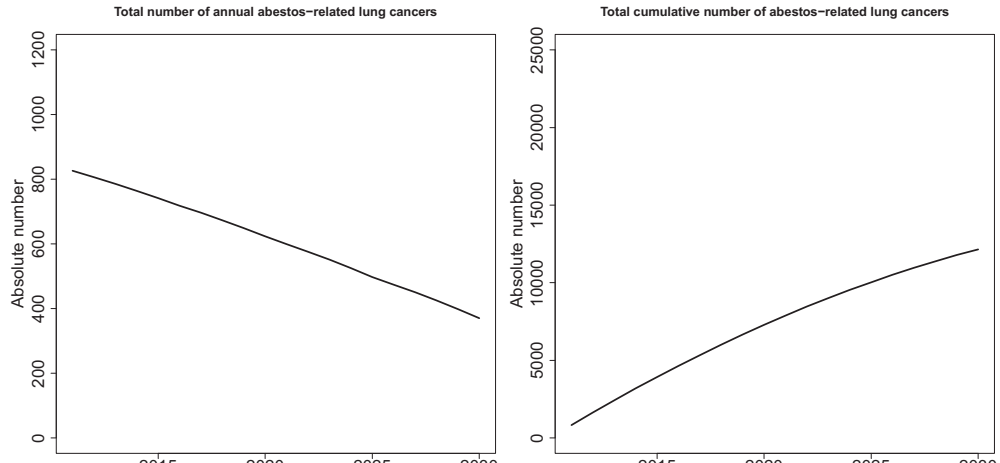


Figure 2: Total number of asbestos-related lung cancer cases in 2011-2030 as estimated by the PAR model; i.e., the population attributable risk model.

Model 3: the life table model

Based on available exposure information it was estimated that in 2010 about 20% of all individuals of 50 years or older had been occupationally exposed to asbestos. Among these individuals assumed exposed the average cumulative exposure was low, ranging from 0.1 to 2.2 f-y/ml in men and from 0.2 to 1.0 f-y/ml in women (table 2).

Initially, the cumulative number of asbestos-related lung cancers between 2011 and 2030 was estimated to be around 1800 in model 3. However, our calibration procedure indicated that either cumulative exposures or mesothelioma risks needed to be about 3.5 times higher in order to obtain in 2011 the same number of mesothelioma cases as was observed on average in 2006-2010. We assumed that it was more likely that we underestimated the cumulative exposure and therefore increased the exposure estimates by a factor 3.5. Assuming this increase is justified, the total number of lung cancers due to asbestos exposure between 2011 and 2030 was estimated to be about 6,300 cancers. Figure 3 has a layout similar to figures 1 and 2, and shows the number of future asbestos-related lung cancers as estimated by the life table Model after recalibration. Here, the annual number of asbestos-related lung cancers decreased consistently over time, from 389 in 2011 to 215 in 2030.

Table 2: Estimated population based asbestos exposure prevalence and cumulative exposures in 2010

	Number of persons ever exposed (%)	Cumulative asbestos exposure level (f-y/ml)
Men		
aged < 20	0	-
aged 20-40	4.2	0.1
aged 40-49	15.6	0.2
aged 50-59	19.7	0.6
aged 60-69	21.1	1.2
aged 70-79	22.3	1.7
aged 80-89	21.5	2.1
aged 90-94	17.0	2.2
Women		
aged < 20	0	-
aged 20-40	1.1	0.2
aged 40-49	1.2	0.4
aged 50-59	1.5	0.4
aged 60-69	1.6	0.5
aged 70-79	1.1	0.6
aged 80-89	1.1	1.0
aged 90-94	1.1	1.0

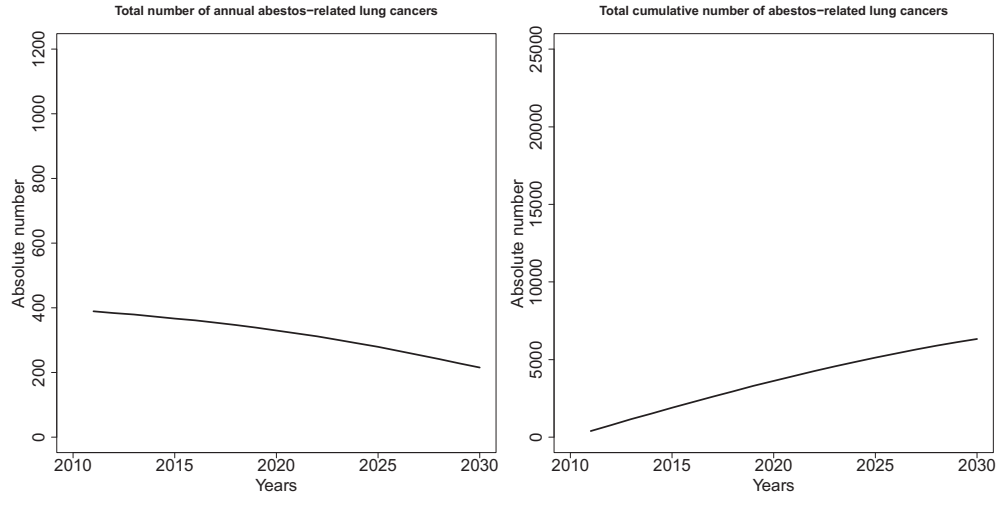


Figure 3: Total number of asbestos-related lung cancer cases in 2011-2030 as estimated by the life table model

Discussion

In this study we compared different model types that may be used to estimate the number of asbestos-related lung cancers instead of applying just a single model. The first model was relatively simple and required little evidence, estimating asbestos-related lung cancer cases directly from observed and predicted mesothelioma cases in an Age-Period-Cohort (APC) analysis. The second model required evidence on the fraction of lung cancer cases attributable to asbestos exposure. The third model was the most comprehensive, requiring actual exposure information and exposure-response functions for mesothelioma and lung cancer to perform a life table analysis on all individuals in the Dutch population. We found substantial differences in model requirements, indicating that it may not be possible to always apply some of the models considered. In addition, our comparison indicated that none of the investigated models is best with respect to all modeling aspects. The APC model is quite specific for this type of disease, whereas the life table model requires evidence that may not be easily available or is relatively uncertain. Moreover, the life table analyses was calibrated based on observed mesothelioma cases making this particular modeling approach specific to this exposure-disease association as well. As a result, the choice for any particular model may be guided by available evidence, the quality of the evidence and the modeling goal, using the comparison shown in table 1.

Given the results of the three models applied, the expected number of asbestos-related lung cancer cases in the Netherlands in the period 2011-2030 varies from 6,300 to 22,150. The highest number of cases was estimated when the number of asbestos-related lung cancers was related to the number of mesothelioma cases (Model 1), whereas the lowest number of cases was estimated by the (calibrated) life table method (Model 3).

It is not straightforward to determine which model is likely to provide the best estimation of the future number of asbestos-related lung cancers. From Table 1 it is apparent that there is no single 'best' model for all settings, rather, each model has its own (dis)advantages. In all models, underlying assumptions and uncertainty in input values substantially decrease robustness of the results. We can, however, identify the main uncertainties for each of the three models.

The main uncertainties in model 1 relates to the choices in the statistical APC modeling techniques, expectations about future developments of birth cohorts risks and the ratio between mesothelioma and asbestos-related lung cancer. Advanced models suggest that the peak of mesothelioma might be earlier with a more rapid decline thereafter than estimated from simple APC models. In these advanced models current mortality is related to past asbestos exposure and do not assume the same age distribution of mortality in different birth cohorts.[26;27] We chose to use the same method as an earlier study in the Netherlands of 12 years ago[28] as their derived predictions appear to closely match actual observations

over the period 2000 to 2011. Compared to that earlier study our future annual number of mesothelioma cases was about 20% higher. However, their predictions compromised only pleural mesothelioma cases in persons up to age 85 years which may have resulted in lower predictions. If we assume that our predictions might be overestimated by maximal 20% our expected number of asbestos-related lung cancers would be lowered by about 5000 cases. The assumption that birth cohorts beyond 1965 had zero risk appeared to be reasonable as the estimated risk for the male birth cohort of 1965 was indeed very low (the APC model estimated for the birth cohort of 1960 a 65% lower risk compared to the birth cohort of 1940). Therefore, our sensitivity analysis in which the birth cohort of 1970 was assigned the risk of the birth cohort of 1965 yielded results very similar to our standard analysis. The greatest uncertainty in model 1 is likely to be in the ratio between mesothelioma and lung cancer, which depends strongly on asbestos fibre type.[12] Different types of asbestos have been used in the past in the Netherlands. For example, asbestos sprays contained amosite whereas asbestos cement products typically contained chrysotyle but may also contained crocidolite. Therefore, it is hard to define a single ratio that can be applied universally to the general Dutch population. Using the ratio previously estimated for chrysotile (a smoking adjusted ratio of 1:3) or amosite (a smoking adjusted ratio of 1: 4.9) instead of mixed fibres would have more than doubled the estimated number of asbestos-related lung cancers. If the ratio for crocidolite had been applied (a smoking adjusted ratio of 1:0.6) estimates would have decreased by more than 50%. One could also believe that a ratio of 1:1 might be more appropriate. This ratio has been observed and used for estimations in the UK.[2;29;30] As the number of mesotheliomas are comparable between the Netherlands and the UK, it has been suggested that these countries are also comparable regarding asbestos exposure. Moreover, in the meta-analyses of McCormack et al. there was a large amount of unexplained between study variability even after stratification on fibre type: the interquartile range of the ratio for studies with mixed asbestos fibres was 1.1 to 4.4. Discrepancies in the ratios might be explained by differences in follow-up time, asbestos industry, (mean) age of individuals at time of exposure, and the level of asbestos exposure.[12] The majority of the studies included in the meta-analyses compromised cohorts of highly exposed individuals from which the ratio between mesthelioma and lung cancer was estimated. However, asbestos exposures at the population level are likely to be lower. Hence, it might be inappropriate to apply the meta-analyzed ratio to the number of mesothelioma cases as observed in the general population. Also, one might expect that the ratio between mesothelioma and occurrence of lung cancer to decrease over time as latency time may be shorter for lung cancer than for mesothelioma.[31]

The main uncertainty in model 2 relates to the applied PAR. We used a PAR of 11.6% estimated from a Dutch case cohort study by van Loon et al.[8] Although this estimate appears to be reasonable compared to other studies, higher and lower PARs have also been reported.[6;7;32-34] A systematic review of asbestos-related cancer in Europe estimated the PAR to be between 5.7 and 19%.[32] Differences in estimates of the PAR are likely to be related to

the amount of asbestos use in the past. In the study by van Loon et al.[8], asbestos exposure was assessed by experts and classified according to *no exposure*, *possible exposure*, *probable exposure* and *near certain exposure*. Although a clear trend in lung cancer risk over these categories was observed the PAR was simplified by using the relative risk for *ever exposed* versus *never exposed* whereas the proportion of ever exposed was estimated by the probability of exposure among ever exposed individuals. As a result the final PAR could be an underestimation. In addition, a PAR estimated 15 years ago may no longer be applicable as other competing lung cancer causes and risks may have changed over time. For example, the numbers of lung cancers due to smoking are likely to be lower as the number of smokers decreased in the last decades.[35] As such the PAR of asbestos might be higher compared to 15 years ago. Due to the uncertainties in the PAR estimate results from these analyses are also relatively uncertain**.**

The main uncertainties in model 3 relate to the exposure-response relationship between asbestos and lung cancer risk as well as to the estimated number of asbestos exposed individuals and the estimated cumulative exposure levels. Our estimates of the number of individuals exposed appear to be relatively high whereas the estimate of the cumulative exposure levels appear to be low. To assess the accuracy of our estimates we determined the number of mesothelioma cases using the life table analysis and compared them to observed mesothelioma cases in 2006-2010. Results indicated that based on our input parameters of risk, exposure prevalence and cumulative exposure levels that we significantly underestimated the number of mesothelioma cases (127 versus 461). This could be the result of either using a lower than actual Km-value or because of underestimation of the prevalence or intensity of exposure. We deemed it more likely that we underestimated the cumulative exposure levels and as such increased the exposure levels by a factor 3.5. There is also uncertainty in which K_L-value to use. A recent meta-regression analyses estimated the K_L-value to be in the range of 0.0033-0.0075.[36] This summary K_L -value was based on mostly high exposed industrial populations and as such may not be directly applicable to the general population. We therefore used a K_L-value of 0.028 which was derived based on a large pooled-analyses using the same exposure approach used in this study.[22] Given the similarities in methods we assumed this K_L-value to be the most appropriate albeit that the K_L-value derived from this study is at the high end of reported K_L values. A high K_L-value has also been observed in a previous population based study.[37]

The advantage of the life table model is that it can incorporate more specific evidence of the relation between asbestos-related lung cancer and asbestos than compared to the other model. For example there is evidence that the effect of asbestos exposure may decline in rates after long latencies.[38] However, this may also add an extra uncertainty to the model. Moreover, a disadvantage is that exposure data on asbestos in the general population is often of limited quality and as such uncertain assumptions have to be made about asbestos exposure in the general population.[3;4]

In conclusion, the preferred method for estimating asbestos-related lung cancer cases in the general population necessarily depends on the available evidence. Robustness of any model depends highly on the quality of evidence. Therefore, a more comprehensive model is not necessarily better than a simple one. However, given the uncertainties in all models it is useful to construct different models, if sufficient evidence is available, and to compare their results. Only by comparing results from different methods insight is gained into the robustness of the estimated number of cases. Results obtained by any one specific method should always be interpreted with caution unless the data collection and analysis is of undeniably high quality. We show that using three different methods results in different absolute estimates of the asbestos-related lung cancers varying by a factor of 3.5. As such the exact impact of asbestos exposure on the lung cancer burden remains uncertain.

Acknowledgments

This study was funded by the Institute for Asbestos Victims, the Netherlands. We would like to thank Dr. P van de Brandt for the use of NLCS data. We also would like to thank Dr. S. Peters and Dr. H. Kromhout for use of DOM-JEM.

Reference List

1. Straif K, brahim-Tallaa L, Baan R, Grosse Y, Secretan B, El GF et al. A review of human carcinogens--part C: metals, arsenic, dusts, and fibres. Lancet Oncol 2009; 10:453-454.
2. Darnton AJ, McElvenny DM, Hodgson JT. Estimating the number of asbestos-related lung cancer deaths in Great Britain from 1980 to 2000. Ann Occup Hyg 2006; 50:29-38.
3. Nicholson WJ, Perkel G, Selikoff IJ. Occupational exposure to asbestos: population at risk and projected mortality--1980-2030. Am J Ind Med 1982; 3:259-311.
4. Walker AM, Loughlin JE, Friedlander ER, Rothman KJ, Dreyer NA. Projections of asbestos-related disease 1980-2009. J Occup Med 1983; 25:409-425.
5. Gasparrini A, Pizzo AM, Gorini G, Seniori CA, Silvestri S, Ciapini C et al. Prediction of mesothelioma and lung cancer in a cohort of asbestos exposed workers. Eur J Epidemiol 2008; 23:541-546.
6. Boffetta P, Autier P, Boniol M, Boyle P, Hill C, Aurengo A et al. An estimate of cancers attributable to occupational exposures in France. J Occup Environ Med 2010; 52:399-406.
7. Gustavsson P, Ahlbom A, Andersson T, Scheele P. Calculation of fractions of lung cancer incidence attributable to occupational exposure to asbestos and combustion products in Stockholm, Sweden. Eur J Epidemiol 2003; 18:937-940.
8. van Loon AJ, Kant IJ, Swaen GM, Goldbohm RA, Kremer AM, van den Brandt PA. Occupational exposure to carcinogens and risk of lung cancer: results from The Netherlands cohort study. Occup Environ Med 1997; 54:817-824.
9. De Vos Irvine H, Lamont DW, Hole DJ, Gillis CR. Asbestos and lung cancer in Glasgow and the west of Scotland. BMJ 1993; 306:1503-1506.
10. Lilienfeld DE, Mandel JS, Coin P, Schuman LM. Projection of asbestos related diseases in the United States, 1985-2009. I. Cancer. Br J Ind Med 1988; 45:283-291.
11. Marinaccio A, Scarselli A, Binazzi A, Mastrantonio M, Ferrante P, Iavicoli S. Magnitude of asbestos-related lung cancer mortality in Italy. Br J Cancer 2008; 99:173-175.
12. McCormack V, Peto J, Byrnes G, Straif K, Boffetta P. Estimating the asbestos-related lung cancer burden from mesothelioma mortality. Br J Cancer 2012; 106:575-584.
13. Carstensen B. Age-period-cohort models for the Lexis diagram. Stat Med 2007; 26:3018-3045.
14. Robertson C, Gandini S, Boyle P. Age-period-cohort models: a comparative study of available methodologies. J Clin Epidemiol 1999; 52:569-583.
15. CBS Statline (database). Voorburg: Centraal Bureau voor de Statistiek (Statistics Netherlands). [electronic resource: http: www.statline.cbs.nl].
16. van der Bij S, Koffijberg H, Burgers JA, Baas P, van d, V, de Mol BA et al. Prognosis and prognostic factors of patients with mesothelioma: a population-based study. Br J Cancer 2012; 107:161-164.
17. Nesti M, Marinaccio A, Gennaro V, Gorini G, Mirabelli D, Mensi C et al. Epidemiologic surveillance for primary prevention of malignant mesothelioma: the Italian experience. Med Lav 2005; 96:338-346.
18. Netherlands cancer registry. Lung cancer and mesothelioma in the Netherlands 1989-1997. 2000. Utrecht: Association of Comprehensive Cancer Centers 2000.
19. Burdorf L, Swuste PH, Heederik D. A history of awareness of asbestos disease and the control of occupational asbestos exposures in The Netherlands. Am J Ind Med 1991; 20:547-555.
20. Virta RL. Mineral commodity profiles-Asbestos: U.S. Geological Survey Circular 1255-KK. 56. 2005.
21. van den Brandt PA, Goldbohm RA, van 't Veer V, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990; 43:285-295.
22. Peters S. Quantitative exposure assessment in community-based studies. 2012. ISBN: 978-90-3935-699-9.
23. Peters S, Vermeulen R, Olsson A, Van GR, Kendzia B, Vincent R et al. Development of an exposure measurement database on five lung carcinogens (ExpoSYN) for quantitative retrospective occupational exposure assessment. Ann Occup Hyg 2012; 56:70-79.
24. Cox D, Oakes D. Analysis of Survival Data. London: Chapman and Hall; 1984.
25. Health council of the Netherlands. Asbest. Risico's van milieu- en beroepsmatige blootstelling. 2010.
26. Hodgson JT, McElvenny DM, Darnton AJ, Price MJ, Peto J. The expected burden of mesothelioma mortality in Great Britain from 2002 to 2050. Br J Cancer 2005; 92:587-593.
27. Tan E, Warren N, Darnton AJ, Hodgson JT. Projection of mesothelioma mortality in Britain using Bayesian methods. Br J Cancer 2010; 103:430-436.

28. Segura O, Burdorf A, Looman C. Update of predictions of mortality from pleural mesothelioma in the Netherlands. Occup Environ Med 2003; 60:50-55.
29. Brown T, Darnton A, Fortunato L, Rushton L. Occupational cancer in Britain. Respiratory cancer sites: larynx, lung and mesothelioma. Br J Cancer 2012; 107 Suppl 1:S56-S70.
30. Rushton L, Hutchings S, Brown T. The burden of cancer at work: estimation as the first step to prevention. Occup Environ Med 2008; 65:789-800.
31. Selikoff IJ, Hammond EC, Seidman H. Latency of asbestos disease among insulation workers in the United States and Canada. Cancer 1980; 46:2736-2740.
32. Albin M, Magnani C, Krstev S, Rapiti E, Shefer I. Asbestos and cancer: An overview of current trends in Europe. Environ Health Perspect 1999; 107 Suppl 2:289-298.
33. Li P, Deng SS, Wang JB, Iwata A, Qiao YL, Dai XB et al. Occupational and environmental cancer incidence and mortality in China. Occup Med (Lond) 2012; 62:281-287.
34. De Matteis S., Consonni D, Lubin JH, Tucker M, Peters S, Vermeulen RC et al. Impact of occupational carcinogens on lung cancer risk in a general population. Int J Epidemiol 2012; 41:711-721.
35. Stivoro. Trendpublicatie precetange rokers. Percentage rokers in de Nederlandse bevolking 1958 - 2011 (in Dutch). Den Haag: STIVORO, Dutch Foundation for Smoking and Health; 2012.
36. van der Bij S, Koffijberg H, Lenters V, Portengen L, Moons K, Heederik D et al. Lung cancer risk at low cumulative asbestos exposure: meta-regression of the exposure-response relationship. submitted. 2012.
37. Gustavsson P, Nyberg F, Pershagen G, Scheele P, Jakobsson R, Plato N. Low-dose exposure to asbestos and lung cancer: dose-response relations and interaction with smoking in a population-based case-referent study in Stockholm, Sweden. Am J Epidemiol 2002; 155:1016-1022.
38. Magnani C, Ferrante D, Barone-Adesi F, Bertolotti M, Todesco A, Mirabelli D et al. Cancer risk after cessation of asbestos exposure: a cohort study of Italian asbestos cement workers. Occup Environ Med 2008; 65:164-170.

Part 3

Methodological considerations

Chapter 7

Dealing with heterogeneity in diagnostic meta-analyses

S van der Bij
P Zuithoff
H Koffijberg
KGM Moons
JB Reitsma

Submitted

Abstract

Background: The bivariate regression model for sensitivity and specificity has become a popular model for meta-analyzing data from diagnostic accuracy studies. Our aim is to generate more insight and to enhance use of the bivariate model in diagnostic test accuracy reviews.

Methods: The key complexity of the bivariate model is the presence of the two possibly correlated outcomes rather than a single outcome as in a traditional random effects meta-analysis. We discuss the interpretation, relevance, and ways of presenting the between-study variances and covariance parameter in the bivariate model.

Results: Many diagnostic reviews focus on the pooled sensitivity and specificity, but additional insight is generated by the between-study variability parameters. Similar to other random effects meta-analysis models, higher levels of between-study variability promote the search for possible explanations because unexplained variability is likely to lower the strength of recommendations of a review. Prediction intervals express the impact of between-study variability on a clinically relevant scale. Because of the possible correlation between the two outcome measures, prediction ellipses are more informative than intervals.

Conclusion: The bivariate model is a valid and useful method to describe sensitivity and specificity and their variability in diagnostic reviews where each study reports a single 2-by-2 table. Reviews currently provide limited information on the amount of variability across included studies. This paper provides guidance how to assess, interpret, and report the impact of between-study variability.

Introduction

Since its description in 2005, the bivariate regression model for sensitivity and specificity [1] has become a popular model for meta-analyzing data from diagnostic accuracy studies. It has since been cited over 230 times <web of science, assessed June 2012> and it is one the methods recommended by the Cochrane collaboration.[2] Because of the two possibly correlated outcomes in diagnostic accuracy studies (i.e. sensitivity and specificity), the bivariate model is inherently more complex than statistical methods used in therapeutic intervention reviews where there is only single outcome measure (for example odds ratio, relative risk, risk difference or standardized mean difference). Consequently, many researchers struggle with understanding, fitting and interpreting the results of the bivariate model. In this paper we will explain the key concepts of the bivariate model for a non-technical audience. Our aim is to generate more insight and to promote better use of the bivariate model in diagnostic test accuracy reviews.

In the first section, we describe the parameters of the basic bivariate model. The second section explains the meaning and relevance of the between-study variances and covariance parameters in the bivariate model. The third section describes the different ways of assessing and informing the reader about between-study heterogeneity. The last sections discuss the merits and disadvantages of other accuracy measures than sensitivity and specificity that can be derived from the bivariate model. We end with some concluding remarks and guidance for presenting results.

The bivariate model of sensitivity and specificity

Diagnostic tests results are frequently reported as negative and positive test results and compared to the reference test in a 2-by-2 table to evaluate how well the test can distinguish between patients with and without the disease of interest. Based on such a 2-by-2 table several measures of accuracy can be calculated. The most popular ones are the test's sensitivity (i.e. the proportion positive test results among diseased) and specificity (i.e. proportion negative test results among those without disease). The bivariate model preserves this two-dimensional nature of the data by analyzing pairs of sensitivities and specificities from individual studies jointly while acknowledging the possible association between them. The bivariate model is a random effects model anticipating that the true (logit transformed) sensitivities and specificities are not the same across studies but vary between studies. Similar to other random effects meta-analyses, the assumption is that the true values of logit sensitivity ($\mu_{A,i}$) and specificity ($\mu_{B,i}$) from individual studies come from approximately normal distributions with certain mean values (μ_A and μ_B) and level of variability around these means (σ_A and σ_B) (figure 1). In addition to the variability in true values, there is an additional source of variation when examining the observed (crude) values of sensitivity and specificity within a review: sampling variation. Observed values in a study differ from their true values due to sampling variation as only a limited number of

patients are included in any particular study. Observed values from larger studies are more likely to be closer to their true values than observed values from smaller studies. This is the rationale that larger studies get more weight when estimating the pooled (or mean) value. Because of the two sources of variation (between-study variability in true values and sampling variability within studies), random effect models are also referred to as hierarchical models. The way data are assumed to be generated in a review is depicted in figure 1.

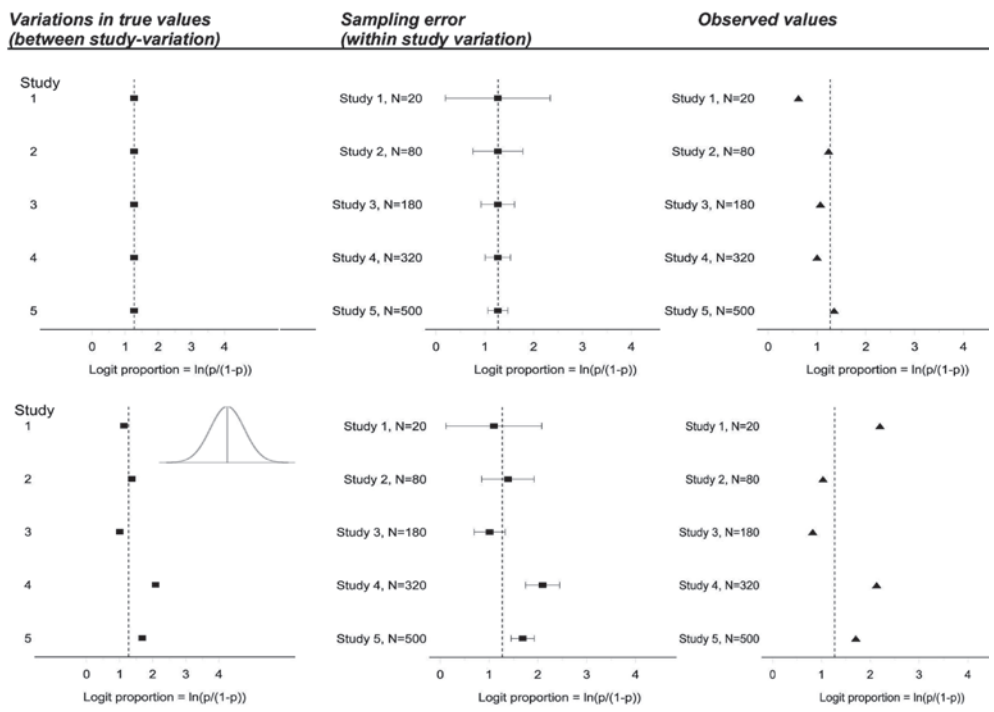


Figure 1: Between-study variance and sampling error in meta-analysis. The first row shows the assumptions of a fixed effect meta-analysis. The second row shows the assumptions of a random effect meta-analysis. In the fixed effect analysis there is no variation in the true values. However, observed values may deviate from their true values due to sampling error. Sampling error decreases with increasing study size. In the random effect analysis there is besides sampling error also variation in the true values.

In technical notation the bivariate model is described by:

$$\begin{pmatrix} \mu_{A,i} \\ \mu_{B,i} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \sum AB\right) \; with \sum AB = \begin{pmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{pmatrix}$$

where μ_A and μ_B are the mean logit-transformed sensitivity and specificity across studies, σ_A^2 and σ_B^2 the between-study variations in logit sensitivity and specificity and σ_{AB} the covariance between logit-transformed sensitivity and specificity. The logit transformation is a popular transformation when analyzing proportions as the range of values now extends from minus to plus infinity thereby avoiding proportions less than 0 and larger than 1 after back transformation.

$$logit\ sensitivity = ln\left(\frac{sensitivity}{1-sensitivity}\right)\ and\ logit\ specificity = ln\left(\frac{specificity}{1-specificity}\right)$$

The meaning of the various parameters of the bivariate model will be described in detail in the next sections. Here, we primarily focus on situations where there is only one 2-by-2 table reported per study.

Basic parameters of the bivariate model

Mean values

The bivariate model produces summary estimates of logit-transformed sensitivity and specificity (μ_A,μ_B) and their 95% confidence interval. These summary values are based on a weighted average of the estimates from the individual studies. As the bivariate model is a random effects model, studies will be weighted in the analyses according to the precision of the estimates within a study and the estimated between-study variation, similar to any random effects meta-analysis. The confidence interval informs the reader about the precision by which the mean or summary values have been estimated.

Between-study variance parameters

The between-study variations in sensitivity (σ_A^2) and specificity (σ_B^2) are similar in definition and meaning to the between-study variation (known as tau, τ) of any random effects meta-analysis. They express the amount of heterogeneity in true values between studies, i.e. the amount of heterogeneity in results between studies that cannot be explained by sampling error given the sample sizes of the studies included. In situations where the between-study variances are (close to) 0, all observed variability can be explained by sampling error (= fixed effect approach) .

Covariance parameter

The complexity in diagnostic test accuracy data is that we have two outcomes (sensitivity and specificity). One approach would be to do a standard meta-analysis twice: one for sensitivity and one for specificity. However, just repeating the analysis means that we assume that the two outcomes are unrelated. In diagnostic accuracy data, however, it is likely that these two outcomes are negatively associated. In other words, if a study has a relatively high sensitivity, it is likely that its specificity will be relatively low. Such a negative association can arise when different studies use different thresholds to decide whether a test produces abnormal (positive) or normal (negative) test results. This is easy to recognize when studies have applied a different threshold for a continuous test result to classify a test result as abnormal (explicit threshold difference), but can also be present if the classification involves a subjective threshold (implicit threshold). An example of an implicit threshold would be the scoring of an image. The bivariate model estimates the amount of association that is present within the data. This is the covariance parameter in the bivariate model. An alternative way

to express this association is by calculating the Pearson's correlation, which ranges from -1 to 1:

$$rho = \frac{\sigma_{AB}}{\sqrt{\sigma_A^2} * \sqrt{\sigma_B^2}}$$

Inclusion of the covariance has the advantage that it leads to a pooled sensitivity and specificity that have a smaller standard error than those arising from two separate (univariate) random effects meta-analyses.[3] Moreover, the correlation provides further insight to describe the relationship between sensitivity and specificity. Figure 2 shows the bivariate distribution of the both logit-transformed and original values of sensitivity and specificity under different simulated scenarios. In scenario (A) a bivariate distribution is displayed where there is no correlation between the two outcomes. Studies vary in their estimates of sensitivity and specificity, but knowing the value of one parameter does not help in predicting what the other value will be (uncorrelated or independent). In the next scenario (B), there is a considerable amount of negative correlation. This has an impact on the shape of the cloud of data points. If sensitivity of a study is high, it is likely that the corresponding specificity will be low. In situations where there is a perfect negative correlation the true values of the studies will be on a straight line. However, this "perfect" straight line may be hard to "eyeball" as the individual data points will deviate from their true underlying values due to sampling error within a study.

In case the two outcomes are related, knowing the value of one parameter will improve the prediction of the other value. The between-study variance parameters describe the total amount of variation in sensitivity or specificity and do not take into account that part of this variation can be explained by knowing the value of the other parameter. In case of correlation, one can estimate the between-study variability in sensitivity at a specific (fixed) value of specificity and vice versa. This is known as the conditional between-study variation:

$$conditional\ \sigma_A^2 = \sigma_A^2 - \frac{\sigma_{AB}^2}{\sigma_B^2}$$

(for estimating the conditional variance in sensitivity at a fixed value of specificity) and

$$conditional\ \sigma_B^2 = \sigma_B^2 - \frac{\sigma_{AB}^2}{\sigma_A^2}$$

(for estimating the conditional variance in specificity at a fixed value of sensitivity)

Knowing the covariance will thus generate further insight in the variability of results between studies.

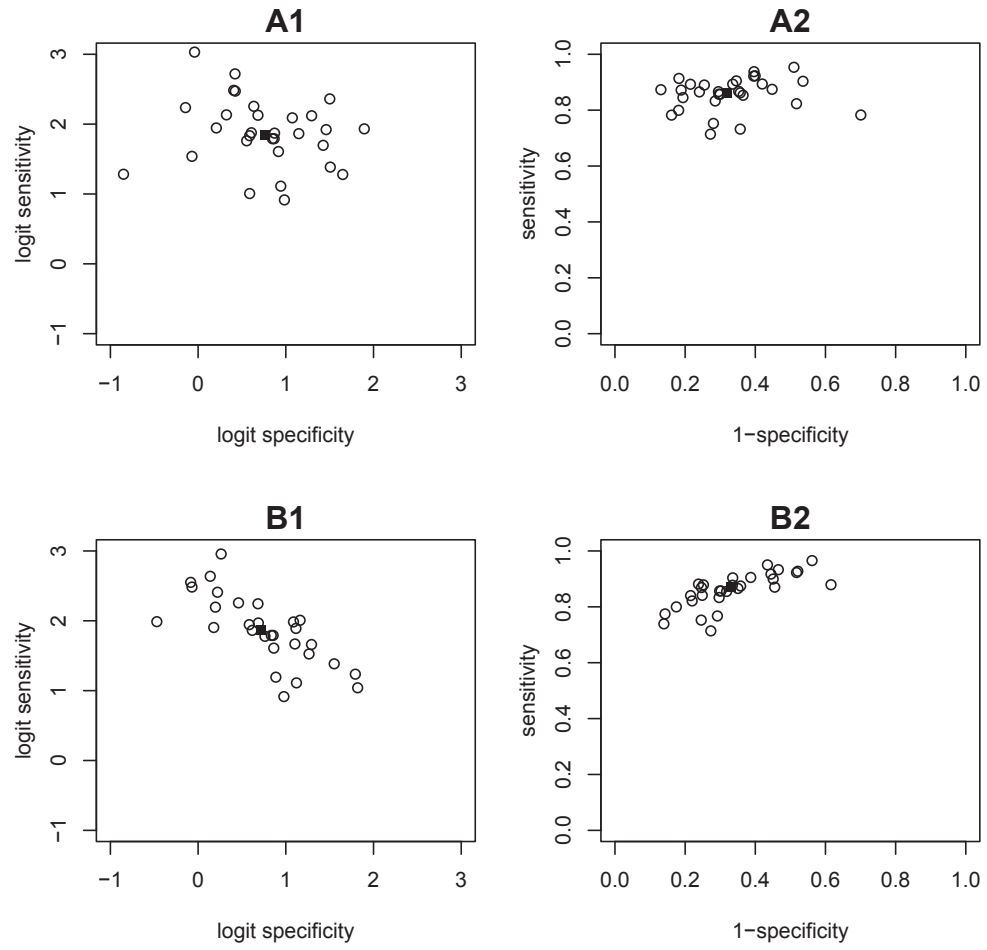


Figure 2: Results from a simulation study where 30 pairs of sensitivity and specificity of individual studies are generated from a normal distribution with a true mean sensitivity of 0.85 and a true mean specificity of 0.70 and with a between-study variance for logit sensitivity of 0.4 and logit specificity of 0.4. In the first row (scenario A) of graphs there is no correlation (covariance) in the second row (scenario B) the correlation is -0.9. The black square presents the bivariate summary estimates of sensitivity and specificity in each panel.

Assessment of heterogeneity

Assessment of heterogeneity in study results is an essential part of meta-analysis. Unexplained heterogeneity reduces the confidence in the overall pooled result because it becomes unclear which value to use or expect in a specific new setting. One might also decide to refrain from pooling estimates when there is a considerable unexplained heterogeneity.

The first step in the assessment of heterogeneity of results across studies is to graphically plot the data. A forest plot is a useful manner to graphically show the variation in estimates of sensitivity and specificity in relation to the precision (width of the confidence interval) of the estimates from individual studies. Additionally, pairs of sensitivity and specificity can be

plotted as points on a receiver operating plot to visualize the heterogeneity and correlation in pairs of sensitivity and specificity.

The next step is to formally examine whether heterogeneity beyond chance exists or to measure the amount of heterogeneity. Statistical testing is problematic because, as with all tests, it is influenced both by the magnitude of the parameter of interest and by sample size. Therefore, it is more insightful to quantify the amount of study heterogeneity rather than to test for it. A frequently used measure to quantify heterogeneity in therapeutic meta-analyses is the I^2 statistic.[4] This statistic does not depend on the number of studies in meta-analyses unlike the well-known Q-statistic.[4] The I^2 statistic describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance). It ranges from 0% (no heterogeneity) to 100% (considerable heterogeneity). However, Rücker et al. demonstrated that this measure is in general of limited use in assessing clinically relevant heterogeneity.[5] In their study, it was shown that the I^2 tends to go to 100% if the precision of the studies included increases. So, a high I^2-statistic might not be due to high between-study variability, but just due to low sampling error within studies (i.e. large studies). Hence, Rucker et al. state that the underlying between-study variability can be best expressed simply by estimating the between-study-variance. In the bivariate model, this is described by σ^2_A for sensitivity and σ^2_B for specificity. However, because of the logit transformation applied when meta-analyzing proportions, the absolute value of the between-study variance is not easily interpretable. Its impact on the original scale (e.g. proportion) depends not only on the absolute value of the between–study variance, but also on the mean value of the proportion. High or low values on the logit scale are more squeeze than values around the value 0 on the logit scale after back transforming to the normal scale. So, points are more widely distributed on the normal scale when a proportion approaches 50% (logit (0.50) = 0), despite similar between-study variability on the logit scale. Hence, a substantial between-study variance appears to have lesser impact on low or high proportions, but more impact on proportions closer to 0.5 (figure 3).

A more informative way to present the amount of between-study heterogeneity is therefore to calculate prediction intervals. A prediction interval shows the range of values where the true value of a proportion coming from a new (comparable) study is likely to be (on the original scale).[6] In other words, if a new large study would be performed comparable to already included studies, there is a 95% probability that its proportion would fall within this interval. Prediction intervals differ from confidence intervals around a summary estimate. For example, a prediction interval will remain relatively large if there is between-study variability within the review, no matter the number of included studies (see table 1), whereas the confidence interval around the pooled value will become smaller when the number of included studies increases.

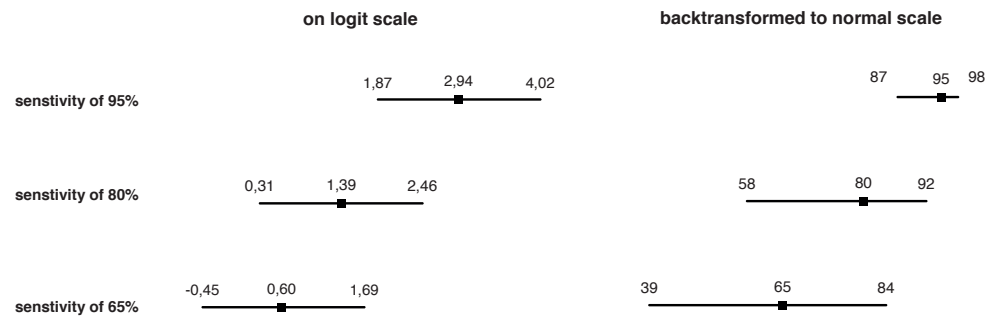


Figure 3: Prediction intervals showing the variability in sensitivity on the logit scale and on the normal scale (after transforming the results of the logit scale to the normal scale). Although the same amount of between-study variation is present on the logit scale for sensitivities of 95%, 80%, and 65%, the impact on the normal scale is different.

Even in situations where there is substantial (unexplained) between-study variation in results, the confidence interval can become very small if sufficient numbers of studies are available (table 1). This difference in behavior can be compared with the difference in behavior between standard error and standard deviation. Standard errors will become smaller if more subjects are included (comparable to confidence intervals), whereas standard deviations describe a feature of the study population which will not become (systematically) smaller if more subjects are examined (comparable to between-study variance).

Table 1: The impact of the amount of between-study variance and number of studies included in the meta-analysis on the confidence interval and prediction interval around a sensitivity of 65%*

Between-study variance	Number of studies included in the meta-analysis N=15		Number of studies included in the meta-analysis N=100	
	confidence interval around mean	prediction interval for a new comparable study	confidence interval around mean	prediction interval for a new comparable study
none (fixed)	0.63-0.68	0.63-0.68	0.64-0.66	0.64-0.66
low	0.62-0.68	0.55-0.75	0.64-0.66	0.55-0.74
high	0.54-0.73	0.27-0.91	0.61-0.68	0.25-0.92

*Simulated number of individuals that underwent the test of interest varied between 70 and 350 among the included studies. Number of patients with the target disease in each study was simulated as 50%.

If substantial heterogeneity is present, prediction intervals provide additional information not captured by confidence intervals.[6] Because of the likely association between sensitivity and specificity, the prediction interval in diagnostic meta-analyses is presented as a prediction ellipse (figure 4).

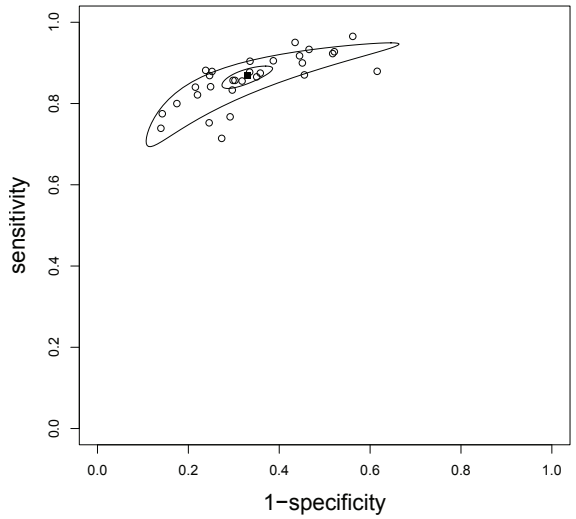


Figure 4: 95% confidence and prediction ellipse. The prediction ellipse (the outer ellipse) shows the range of likely values where the true values of sensitivity and specificity can be expected coming from a new, (comparable) study. The 95% confidence ellipse (the inner ellipse) shows the range of likely values for the summary values of sensitivity and specificity. The black square presents the bivariate summary estimate of sensitivity and specificity. Ellipses were estimated from the simulated data of scenario B in figure 2.

Potential sources of heterogeneity

In case of meaningful between-study heterogeneity a key step is to explore potential sources of heterogeneity if a sufficient number of studies is available. Explaining the differences in test accuracy across studies could lead to a better scientific understanding of the data and more clinically useful conclusions on which to base decisions about medical tests. Typical sources of heterogeneity are clinical or methodological differences among the studies.
Methodological issues are often related to the potential of bias. Assessment of the risk of bias is part of the quality assessment of studies with QUADAS being the most popular instrument for diagnostic accuracy studies.[7;8] To limit heterogeneity due to bias it is recommended to include only studies of high quality. Clinical heterogeneity results from differing patient populations. Factors related to heterogeneity can be included in the bivariate model as a co-variable. Identifying subgroup of studies with more similar results based on a common factor will lower the amount of unexplained between-study variance.

One special source of heterogeneity is variation in the (implicit) threshold between studies of the same test to indicate test positivity. A high negative correlation between sensitivity and specificity might be explained by threshold differences in test positivity. In such situations one can estimate the conditional between-study variation (as previously described) to give an indication of the study heterogeneity that is not related to threshold. However, there are several other factors that can lead to opposing changes in sensitivity and specificity, such as partial verification bias and different patient selections. Hence, if correlation is high,

threshold differences should be regarded as just one of the possible explanations of study heterogeneity.

In meta-analyses where studies applied a different explicit threshold, variability in results can be reduced by selecting sensitivities and specificities that are reported for a common threshold. If one is interested to know how sensitivity and specificity trade-off with each other as threshold varies different approaches can be used where per study only one estimated pair of sensitivity and specificity (corresponding to different diagnostic tresholds) is available. One could stratify results to meta-analyze sensitivity and specificity at different cut-offs or analyzing threshold as a co-variable. If there is little consistency in the applied threshold one could just describe the data at varying thresholds. To really examine the impact of threshold on accuracy, several 2*2 tables from each study are needed or the individual patient data with the continuous test results.[9;10] If just a minority of studies report several 2*2 tables, these extra data points can be added to the analyses when results are stratified according to cut-off.

Other accuracy measures than sensitivity and specificity

The parameters of the bivariate model can be used to produce a summary Receiver Operating Characteristic (sROC) curve. This sROC is estimated by summarizing the bivariate distribution by a single straight line. However its interpretation is not similar to a normal ROC curve in diagnostic studies. Arends et al. showed that five different sROC curves can be derived from the same bivariate distribution [11] One of these curves is identical to the Hierarchical Receiver Operating (HROC) curve of Rutter and Gatsonis. Each of the different sROC curves has its own interpretation and properties. Without additional data, it is impossible to determine which of these five curves would be a better approximation of the true ROC curve that arises when the threshold is varied. To estimate the true average ROC curve one should have multiple two-by-two tables from each study or individual patient data with the continuous results of the test.[1;9] Moreover, drawing an sROC in situations where there is one 2*2 table reported per study suggests that the association between sensitivity and specificity is entirely due to threshold differences which may not be correct for reasons as stated before. Also, if an implicit or one common threshold exists, the main interest is to estimate the average operating point and describe variation rather than to specify an sROC curve as variation due to an implicit threshold might always be present in practice.

Other measures that can be derived from the bivariate model are the diagnostic odds ratio (DOR) and the likelihood ratio.[12;13] These measures can be calculated directly from the bivariate model as they are a combination of the mean sensitivity and specificity. The diagnostic odds ratio (DOR) describes how many times higher the odds are of obtaining a test positive result in a diseased than a non-diseased person and is specified by:

$$DOR = e^{(\mu A + \mu B)}$$

The odds ratio does not provide extra information beyond sensitivity and specificity and therefore less relevant to use.

The likelihood ratio summarizes how many times more (or less) likely patients with the disease are to have a particular result than patients without the disease The larger the positive likelihood ratio is, the greater the likelihood of disease; the smaller the negative likelihood ratio, the lesser the likelihood of disease. The likelihood ratios for positive and negative test results are defined by:

$$LR+= \frac{sens}{1-spec} = \frac{e^{\mu_A}/(1+e^{\mu_A})}{1-\left\{e^{\mu_B}/(1+e^{\mu_B}\right\}}$$

$$LR-= \frac{1-sens}{spec} = \frac{1-\left\{e^{\mu_A}/(1+e^{\mu_A}\right\}}{e^{\mu_B}/(1+e^{\mu_B})}$$

Standard errors and 95% confidence intervals of DOR and likelihood ratios may be derived from the standard errors of sensitivity and specificity using Taylor approximations.[14]

Implication of test results

In clinical practice it is essential to know how to interpret the result of a particular test. Sensitivities and specificities do not provide this information directly. After selecting a prior probability (usually the prevalence), predictive values can be derived to give probabilities of disease for a particular test result based on Bayes theorem.[15] Estimates of the prevalence can be obtained from literature, the weighted mean or median prevalence from the included studies in the review, from a study in a specific setting, or based on opinions from clinical experts. The prior probability can also be defined as the post-test probability resulting from one or more preceding tests. A summary 2-by-2 table can be constructed based on the summary estimates of sensitivity and specificity and a typical value of prevalence to show the effects of performing a test in absolute numbers (table 2). Such a table provides information about the number of false negative and false positive results. This can be done for a range of values of the prevalence and possible alternative combinations of sensitivity and specificity to show the robustness of the results.

Concluding remarks and guidance

In conclusion, the bivariate model is a valid and flexible model for meta-analyzing pairs of sensitivity and specificity. Reviews often focus on reporting mean values for sensitivity and specificity without informing the reader about the amount of unexplained between-study variability. Prediction ellipses are a clinically relevant way to demonstrate variability in reported accuracy of diagnostic tests in a systematic review. Translating summary results of sensitivity and specificity to absolute numbers of patients with true positive, false-positive,

true negative and false-negative test results is a meaningful way to show the clinical consequences of using a particular test.

Table 2: Clinical implication of applying a test with a mean sensitivity of 85% and a mean specificity of 75% in a population of 1000 persons with a disease prevalence of 10%

	Disease prevalence=10% N=1000		
Test result :	**Target disease present N=100**	**Target disease absent N=900**	**Posterior probability**
Abnormal	85	225	PPV=0.27
Normal	15	675	NPV=0.98
	sensitivity=0.85	specificity=0.75	

PPV=Positive predictive value; NPV= Negative predictive value

Reference List

1. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58:982-990.
2. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from: http://srdta.cochrane.org/. 2012.
3. Riley RD, Abrams KR, Lambert PC, Sutton AJ, Thompson JR. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. Stat Med 2007; 26:78-97.
4. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327:557-560.
5. Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008; 8:79.
6. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011; 342:d549.
7. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003; 3:25.
8. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155:529-536.
9. Hamza TH, Arends LR, van Houwelingen HC, Stijnen T. Multivariate random effects meta-analysis of diagnostic tests with multiple thresholds. BMC Med Res Methodol 2009; 9:73.
10. Hollevoet K, Reitsma JB, Creaney J, Grigoriu BD, Robinson BW, Scherpereel A et al. Serum mesothelin for diagnosing malignant pleural mesothelioma: an individual patient data meta-analysis. J Clin Oncol 2012; 30:1541-1549.
11. Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MG, Stijnen T. Bivariate random effects meta-analysis of ROC curves. Med Decis Making 2008; 28:621-638.
12. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol 2003; 56:1129-1135.
13. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. Stat Med 2008; 27:687-697.
14. Oehlert GW. note on the delta method. American Statistician 46(1). 1992. Ref Type: Generic
15. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ 1994; 309:102.

Chapter 8

General Discussion

General discussion

Asbestos has been used worldwide, and is still used in developing countries because of its strength and resistance to heat. However, the negative health-related consequences of its usage have become clearly apparent and widely recognized. This thesis focused on the most frequent asbestos induced cancers: mesothelioma and lung cancer.

Main conclusions based on articles of this thesis

Part 1: mesothelioma

Many countries have a financial arrangement to handle claims of patients with mesothelioma. These patients are often only entitled to financial compensation if the diagnosis is pathologically confirmed. In the Dutch situation, as described in chapter 2, in about 6% of the patients with suspected malignant mesothelioma pathologic material is not available or insufficient for diagnosis. For patients without pathologic diagnosis clinical assessment based on other diagnostic information could be an option. Within the Dutch system (chapter 2) a diagnosis of malignant mesothelioma was established in 80% of these patients based on judgment of clinical experts reviewing all available clinical and radiological data. Here, the structured and protocol based judgment procedure may support the assessment by qualified experts (chapter 2)

To increase the chance of a pathologic diagnosis in patients suspected with mesothelioma it is recommended to obtain biopsy material. Therefore, it is important to refer patients early for diagnostic procedures. However, the diagnosis of mesothelioma is not straightforward and symptoms are usually non-specific. Obtaining biopsy material is an invasive procedure and is commonly performed only if there are acceptable and guideline supported indications. To facilitate a non-invasive diagnostic work-up it would be valuable to have non-invasive diagnostic procedures that accurately confirm or exclude the diagnosis of mesothelioma. Accordingly, as shown in chapter 3, numerous non-invasive marker tests have emerged, based on the increasing understanding of the molecular and biological pathways of mesothelioma. These include many immunohistochemical markers for cytological analysis and serum markers. Our results showed that the most frequently studied immunohistochemical markers for cytological analysis were EMA, Ber-Ep4, CEA, and calretinin. The most frequently investigated serum marker was SMRP. However, the evidence to properly assess the value of these non-invasive markers in the diagnosis of mesothelioma is currently very limited. Our systematic review investigated all available studies on these markers and showed that these publications were considerably heterogeneous in terms of investigated patients, type of study designs (e.g. type of markers, outcomes, and selection of study subjects) and reported statistical results. Nevertheless, our analyses indicated that CEA, Ber-EP4 and calretinin were most valuable in discriminating mesothelioma from other malignant diseases. EMA and SMRP were most valuable in discriminating mesothelioma from non-malignant diseases.

None of the markers performed well to differentiate mesothelioma from both malignant and non-malignant diseases. Furthermore, all the immunohistochemical markers, especially CEA, were mainly of value in exclusion of mesothelioma as sensitivity (and thereby the negative predictive value) was in general high. The specificity of these markers varied and depended on the comparison group and therefore the differential diagnosis. Only SMRP at a high cut-off value was a strong indicator for mesothelioma but the associated poor sensitivity of SMRP (<60%) clearly limits its added value to be used in practice.

There is a particular interest in a valuable serum marker specific for mesothelioma as it might be applied by the patient's physician at start of the first symptoms and therefore especially may aid earlier diagnosis of mesothelioma. Generally, a rapid diagnosis will also have implications on survival by improving the effect of treatment. To date, the survival of malignant mesothelioma patients remains poor. One year after diagnosis we observed that only 47% of the patients were still alive (chapter 4). Predictors strongly associated with survival were age, mesothelioma localization and subtype. Our results are consistent with other population-based studies.[1-4] Hence, the prognosis is still far from satisfactory for the majority of patients even though more treatment options have become available.

Part 2: asbestos-related lung cancer risk

Previous estimated lung cancer risks per unit of asbestos exposure are mainly based on, and applicable to, highly exposed industrial cohorts. To determine the risk of asbestos-related lung cancer in the general population we assessed the risk at low cumulative asbestos exposure estimates based on all available literature (chapter 5). Based on our meta-analyses, the increase in relative risk of lung cancer due to low asbestos exposure may be larger than expected from previous studies.[5;6] This suggests that, in general, a larger fraction of lung cancer incidence may be attributable to, relatively low, asbestos exposure levels than previously estimated.

Given the shift from relatively few individuals with high exposure to relatively many individuals with low exposure, in developed nations with strict regulations and bans on asbestos use, our results may have substantial implications for the number of future lung cancers attributable to asbestos exposure.

The question regarding the exact number of expected asbestos-related lung cancers over time remains however difficult to answer given the inherent uncertainty in input parameters. To obtain an estimate of the expected future number of lung cancers attributable to asbestos in the general population we compared three methods which differ in complexity and required evidence (chapter 6). Application of these three methods to the Netherlands provided estimates of the total number of lung cancers attributable to asbestos ranging from 6,300 to 27,200 in the period 2011-2030. We concluded that the preferred method for estimating asbestos-related lung cancer cases in the general population necessarily depends on the available evidence and robustness of any model depends highly on the

quality of evidence. Therefore, a more comprehensive model is not necessarily better than a simple one. Furthermore, only by comparing results from different methods insight is gained into the robustness of the estimated number of cases. Results obtained by any one specific method should therefore always be interpreted with caution unless the data collection and analysis is of undeniably high quality.

Part 3: methodological considerations

Based on experience from our own systematic review on the mesothelioma diagnostic markers (chapter 3), the last part of this thesis (chapter 7) involved a methodological exercise on how to address issues arising from heterogeneity in diagnostic accuracy studies when conducting a systematic review of such studies. When in a diagnostic meta-analysis, different primary studies of the same continuous test (such as the mesothelioma markers described in chapter 3) report the explicit cut-off value used to estimate the marker's or test's sensitivity and specificity, several options (such as summary ROC curves) are available to describe heterogeneity in the estimated sensitivities and specificities due to variation in the used cut-off values. Summary ROC curves, however, are of limited value when per study only one cut-off value of a continuous test (e.g. mesothelioma marker) has been reported. In such circumstances, and in situations where no explicit cut-off value for a tests can be defined anyhow (as often is the case for imaging tests), so-called prediction intervals are a clinically relevant approach to demonstrate variability in reported accuracy of diagnostic tests in a systematic review.

Implications for current and future policy making, practice and research

Diagnosis of Mesothelioma

In November 1998, an agreement was successfully negotiated in the Netherlands between the ministry of justice, the ministry of social security, the trade unions and asbestos victims to ensure that future claims would be handled in a swift and socially acceptable manner. For that purpose the Institute for Asbestos Victims was founded. Its primary task is to support mesothelioma patients in this fast-reward claim process, to provide cash advances, and to organize a short track protocol to select patients who are entitled to additional compensation. In these patients, it is presumed that the previous employer or his/her insurance company are liable or at least have a legal responsibility to deal with the claim. The Institute for Asbestos Victims collects the evidence on behalf of the patient in order to establish the liability. So either for an advance compensation irrespective of the nature of the asbestos exposure or in order to obtain full compensation based on employer's liability, it is therefore essential to obtain a definite diagnosis of malignant mesothelioma. If an individual applies, the diagnosis of malignant mesothelioma first needs to be confirmed by a 'national panel of pathologists', using both histological and cytological samples. If a diagnosis of malignant mesothelioma cannot be made on the basis of a cytological or

histological evaluation (for whatever reason), subsequently a panel of 'clinical experts' evaluates all available clinical and radiological data, to ultimately determine whether the presence of malignant mesothelioma is more likely than some other diagnosis.
For our studies on the diagnosis of the (potential) mesothelioma cases, data were available from the Institute for Asbestos Victims. Although it took efforts to obtain access to this data source, as the collection of these data was not a-priori intended for scientific research, we showed that the current approach for confirmation of the diagnosis of malignant mesothelioma is a solid process (chapter 2). Naturally, continuous and transparent quality control of the panels' performance remains mandatory. Although not openly expressed, the expert panels act as a mechanism for synthesizing information from a range of sources, drawing on a range of viewpoints, in order to arrive an overall opinion on the presence or absence of asbestos-related diseases to be followed by other professionals. Apart from the effectiveness of this process and the current compensation system, however, it is important that the panels use the latest evidence regarding the tools used for establishing the final diagnosis.

To date, there is no single test available that can diagnose malignant mesothelioma with sufficient accuracy. Even a pathological diagnosis is based on multiple immunohistochemical marker tests and knowledge of prior clinical tests. Still, a final diagnosis can be made in almost all patients if sufficient biopsy material is available. Many markers have emerged that could facilitate a non-invasive diagnosis. Several markers have already entered the market and are used in clinical practice. However, their accuracy and cost effectiveness are not yet systematically investigated and studies containing sufficient numbers of patients are lacking. It is not surprising that several tests therefore disappeared after initial promising results. To optimize current strategies based on biomarkers in serum either for diagnosis or treatment indication, prospective studies in suspected mesothelioma patients are needed. These prospective studies are indeed hard to perform on a single institution basis as the disease under study has a low incidence. Hence, researchers and physicians should join forces to collaborate within multicenter studies to enhance the proper quantification of the diagnostic accuracy of the most promising markers for mesothelioma. Considering the fact that this approach is widely accepted within oncology research, one may wonder why such an infrastructure for mesothelioma research is still not in effect.
As the diagnosis of mesothelioma remains a multifactorial process, objective expert panels are of great value to confirm the diagnosis which, in turn, allows for financial compensation. Although it is recommended to obtain biopsy material to increase the chance of a pathological diagnosis in patients suspected with mesothelioma, it remains ethically questionable to demand an intervention in order to establish the entitlement to financial compensation. If pathologic material is not available or insufficient clinical assessment is an option. It should be further elucidated exactly how well clinical assessment by experts can identify patients with malignant mesothelioma.

Moreover, the process to confirm the diagnosis needs to be as efficient as possible and has to follow a short track as we showed that the prognosis is still very poor (chapter 4). To improve this process future research should focus on the optimal timing of consultation, within the diagnostic process, of a pathologic expert panel by local hospitals.

To speed up the initial diagnostic process, physicians may collect a complete occupational history on possible exposure to asbestos in the workplace or in environmental settings as a first step in linking symptoms to asbestos-related diseases. Additional training in occupational medicine can make medical specialists more aware of potentially relevant asbestos exposure. Moreover, the accurate collection of a complete occupational history in patients with mesothelioma might identify (new) occupations in which workers were at risk of asbestos exposure. To inform persons of possible historical asbestos exposure, certain countries try to identify former workers exposed to asbestos through specific programs.[7-10] These programs are either focused on individuals who worked in specific companies or industries, or cover the whole population of exposed workers. Patients with identified high exposure are then offered a medical surveillance program. An additional benefit of such programs is that further insight might be gained in the long-term effects of asbestos exposure. A survey in France showed that their medical surveillance program induced benefits in terms of detection, information collection and medical surveillance of exposed workers and that participants were highly satisfied with the program.[9] Further benefits of such programs in terms of increased life expectancy and improved treatment effects still needs to be investigated in future studies.

Financial compensation for mesothelioma based on prognosis and suffering.

The results of this thesis may also impact the financial compensation system. Although this thesis was not about financial compensation itself, such a system depends on the numbers of diseased patients, accuracy of the diagnosis, prognosis of disease after diagnosis, and the contribution of asbestos to the disease development.
Amongst other conclusions, the data in this thesis confirmed that life expectancy - independently of quality of life - after the clinical diagnosis, is less than two years for the majority of patients. In the past, before the Dutch mesothelioma agreement was made, a considerable number of patients could therefore not benefit from the compensation. We observed a rather high age in patients with mesothelioma; about 40% of the patients was 70-80 years, and 8% was 80 years or older. The high age in our cohort likely reflects the long latency time between asbestos exposure and diagnosis of malignant mesothelioma. Our results showed an average latency time of 49 years between initial asbestos exposure and diagnosis. As handling of asbestos declined gradually after the 1970s and dropped severely in the 1990s we expect that on average the age at diagnosis will increase in the coming decades. Therefore, future agreements on financial compensation might consider individualized compensation based on life years lost instead of a fixed general compensation for all patients regardless of age at diagnosis.

Financial compensation for asbestos-related lung cancer.

Besides mesothelioma, asbestos is considered to contribute to the development of other malignant diseases, of which lung cancer is the most common. Some European countries like Denmark and Germany have also created financial compensation arrangements for asbestos-related lung cancer. The legal basis for such an arrangement is politically based and very much related to the nation's system on how to compensate occupationally induced disability in general. In the Netherlands, no such arrangement exists for patients with asbestos-related lung cancer. The legal principal is that somebody – a company, a city council, an employer - is to be held liable. This requires at least established causation between asbestos exposure and the occurrence of the lung cancer and the damages experienced by the patient.

From the life table method we estimated that there will be about 41,000 lung cancer patients that have been exposed to asbestos of which about one fourth might be due to asbestos in the period 2011 and 2030. Designing a system to fairly compensate these patients with asbestos-related lung cancer is far more challenging than compensating malignant mesothelioma patients. The actual lung cancer cause cannot be established after diagnosis because the different causes of lung cancer (e.g. smoking) do not lead to very different pathology. It may be possible, though, to estimate the probability that the lung cancer originated from a specific cause, such as smoking or exposure to asbestos.[11] The exposure-response relationship presented in chapter 6 can be used for such a calculation. Then, e.g., using the Helsinki criteria, patients with at least 50% chance that lung cancer is caused by asbestos, may be compensated.[12] It should be further assessed however how accurate asbestos exposure can be (quantitatively) determined in individuals if one wants to fairly compensate asbestos-related lung cancers based on such estimated probabilities. Moreover, little is known about the exact health impact of asbestos-related lung cancer in individuals (such as life years lost) and this also should be further investigated.

Risks of asbestos exposure in the near future

The effect of the regulations introduced to protect workers from asbestos in 1980 and 1990 is becoming more evident in the Western world. Several studies have already shown very low risks of mesothelioma in cohorts of 1960 and beyond,[13;14] which we also observed in our APC model in chapter 6. A trend towards longer latencies and older age in patients with mesothelioma also suggests that the protective regulations have worked. Hence, in the near future a larger part of the patients with mesothelioma will not be related to occupational asbestos exposure anymore, but to background levels of asbestos. The same will also be true for asbestos-related lung cancer. Although some argue that a threshold effect might exist for mesothelioma and asbestos-related lung cancer, the broad opinion is that there is no level of exposure to asbestos below which there is no risk. If any threshold exists it seems to be extremely low and cannot easily be quantified with the current evidence.[12] Our results

in chapter 5 show that lung cancer risks at low asbestos levels may be higher than expected from previous meta-analyses. Moreover, a high risk of lung cancer at low asbestos exposure has also been observed in a population based study.[15] Since asbestos is still in place in many buildings, the hazards of asbestos are not over yet, for instance during remodeling or demolition of such buildings. The question remains whether it is more dangerous to remove asbestos from buildings or to keep asbestos in place. If we keep asbestos in place we will pass the legacy of asbestos on to our grandchildren. If we remove it, we put individuals who remove asbestos at risk. A study has shown that asbestos removal workers have a significant higher risk of asbestos-related cancers than the general population.[16] It is unknown how well safety gear, such as respirators, and other protective regulations work to protect removal workers. Moreover, there is limited evidence of factors that may affect the risk of asbestos at low levels, such as smoking. Hence, it is important to further study the risk of asbestos at current low levels. As the effect of asbestos is an international issue, countries should join forces to collect data for further investigation. By comparing results from studies which vary in measures used to determine asbestos exposure, further insight will be gained regarding the estimated number of asbestos-related cancers and asbestos associated risks. As collection of prospective data may be limited due to time constraints insight could also be gained from other research designs, such as, retrospective analysis of registry data or simulation studies. To adequately protect removal workers and the general population from (low) exposures strict regulations should be enforced.

In developing countries the use of asbestos at present is higher than the use in Western Europe and North America in the 1960s.[17;18] Canada is still mining and exporting chrysotyle in large quantities to China and India as the asbestos industry claims that this particular fibre can be safely used. Although estimates by fibre type show a higher lung cancer risk associated with exposure to amphiboles compared to chrysotile our observed potency differences between different fibre types were lower than the generally held consensus and strongly influenced by a few specific studies. Hence, the effect of chrysotyle should not be underestimated and given the large uncertainties around the risk of chrysotyle exposure should be minimized. The use of asbestos and its risk in developing countries presents a major challenge to the international scientific community in the coming decades.

Reference List

1. Chapman A, Mulrennan S, Ladd B, Muers MF. Population based epidemiology and prognosis of mesothelioma in Leeds, UK. Thorax 2008; 63:435-439.
2. Milano MT, Zhang H. Malignant pleural mesothelioma: a population-based study of survival. J Thorac Oncol 2010; 5:1841-1848.
3. Mirabelli D, Roberti S, Gangemi M, Rosato R, Ricceri F, Merler E et al. Survival of peritoneal malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:194-200.
4. Montanaro F, Rosato R, Gangemi M, Roberti S, Ricceri F, Merler E et al. Survival of pleural malignant mesothelioma in Italy: a population-based study. Int J Cancer 2009; 124:201-207.
5. Lash TL, Crouch EA, Green LC. A meta-analysis of the relation between cumulative exposure to asbestos and relative risk of lung cancer. Occup Environ Med 1997; 54:254-263.
6. Lenters V, Vermeulen R, Dogger S, Stayner L, Portengen L, Burdorf A et al. A Meta-Analysis of Asbestos and Lung Cancer: Is Better Quality Exposure Assessment Associated with Steeper Slopes of the Exposure-Response Relationships? Environ Health Perspect 2011.
7. Huuskonen MS, Koskinen K, Tossavainen A, Karjalainen A, Rinne JP, Rantanen J. Finnish Institute of Occupational Health Asbestos Program 1987-1992. Am J Ind Med 1995; 28:123-142.
8. Bohnker BK, Betts LS, Sack DM, Craft N. Navy Asbestos Medical Surveillance Program, 1995-1999: demographic characteristics and smoking status. Mil Med 2001; 166:966-971.
9. Carton M, Bonnaud S, Nachtigal M, Serrano A, Carole C, Bonenfant S et al. Post-retirement surveillance of workers exposed to asbestos or wood dust: first results of the French national SPIRALE Program. Epidemiol Prev 2011; 35:315-323.
10. Marinaccio A, Branchi C, Massari S, Scarselli A. National epidemiologic surveillance systems of asbestos-related disease and the exposed workers register. Med Lav 2006; 97:482-487.
11. McElduff P, Attia J, Ewald B, Cockburn J, Heller R. Estimating the contribution of individual risk factors to disease in a person with more than one risk factor. J Clin Epidemiol 2002; 55:588-592.
12. Henderson DW, Rodelsperger K, Woitowitz HJ, Leigh J. After Helsinki: a multidisciplinary review of the relationship between asbestos exposure and lung cancer, with emphasis on studies published during 1997-2004. Pathology 2004; 36:517-550.
13. Clements M, Berry G, Shi J, Ware S, Yates D, Johnson A. Projected mesothelioma incidence in men in New South Wales. Occup Environ Med 2007; 64:747-752.
14. Tan E, Warren N, Darnton AJ, Hodgson JT. Projection of mesothelioma mortality in Britain using Bayesian methods. Br J Cancer 2010; 103:430-436.
15. Gustavsson P, Nyberg F, Pershagen G, Scheele P, Jakobsson R, Plato N. Low-dose exposure to asbestos and lung cancer: dose-response relations and interaction with smoking in a population-based case-referent study in Stockholm, Sweden. Am J Epidemiol 2002; 155:1016-1022.
16. Frost G, Harding AH, Darnton A, McElvenny D, Morgan D. Occupational exposure to asbestos and mortality among asbestos removal workers: a Poisson regression analysis. Br J Cancer 2008; 99:822-829.
17. USGS (U.S.Geological Survey). Mineral Industry Surveys: World Asbestos Consumption from 2003 through 2007. Reston, VA: U.S. Geological Survey. 2009.
18. Virta RL. Mineral commodity profiles-Asbestos: U.S. Geological Survey Circular 1255-KK. 56. 2005.

Summary

Summary

Already since 1930 it is known that asbestos is an extremely dangerous fibre. However, since that time, it has taken many years to enforce a complete ban of asbestos in much of the Western world. Moreover, in developing countries the use of asbestos is rapidly increasing.

To date, the impact of the use of asbestos in the Western world is still noticeable. Asbestos is still in place in many buildings and exposure to asbestos may occur during remodeling or demolition of such buildings. Moreover, the health effects of asbestos mostly become apparent after a long period of time as asbestos is related to cancers that may occur many years after exposure. The most notably asbestos-related cancers are malignant mesothelioma and lung cancer. It is expected that the mesothelioma mortality rates continue to rise in the next years.

The *first part* of this thesis relates to diagnostic and prognostic factors in patients with malignant mesothelioma. Although a confirmed diagnosis of malignant mesothelioma is important to ensure proper medical care it is also required to initiate a claim for financial compensation. In the Netherlands, patients with apparent malignant mesothelioma can apply to the Dutch Institute for Asbestos Victims for financial compensation. For each applicant the diagnosis of malignant mesothelioma needs to be confirmed by a 'national panel of pathologists' using both histological and cytological samples. If a diagnosis of malignant mesothelioma cannot be established based on a cytological or histological evaluation a panel of clinical experts evaluates all available clinical and radiological data. This panel of 'clinical experts' then determines whether the presence of malignant mesothelioma is the most likely diagnosis. Accordingly, patients with a pathologically or clinically confirmed diagnosis are candidates for a financial reimbursement in the Netherlands.

To evaluate the diagnostic work-up of patients who applied for financial compensation to the Dutch Institute for Asbestos Victims we determined in chapter 2 how often a pathological or clinical diagnosis could be made and which factors were associated with a final diagnosis of malignant mesothelioma.

We found that in 97 of the 1,498 patients (6%) that applied to the Dutch institute between 2005 and 2008 no pathologic diagnosis could be established. Reasons that patients did not receive a pathologic diagnosis were 1) an uncertain diagnosis (N=54), 2) inadequate tumor samples (N=22), or 3) unavailable tumor samples (N=21). A final pathological diagnosis of malignant mesothelioma could more often be made when biopsy samples were available compared to when only cytological material was available. In the group of patients who did not have a pathologic diagnosis but underwent clinically assessment 80% were considered to have a diagnosis of malignant mesothelioma. None of the clinical features analyzed were strongly associated with a confirmed diagnosis of malignant mesothelioma. Consequently, when patients are only entitled compensation if their diagnosis is pathologically confirmed a small group of patients may not have the possibility to get compensated during their life.

If pathologic material is not sufficient clinical assessment could be an option especially if re-biopsy is not. Clinical assessment is a multivariable process and therefore should be performed carefully by several experts.
Moreover, to increase the chance of a confirmed pathologic diagnosis it is sensible to obtain biopsy material from patients. Therefore, it is important to refer patients early for diagnostic procedures.

To speed up the diagnostic process numerous markers have been evaluated to facilitate the non-invasive diagnostic work-up of malignant mesothelioma. In chapter 3, we conducted a structured review of the diagnostic performance of non-invasive marker tests for the detection of malignant mesothelioma.
We extracted studies on the diagnostic accuracy of serum and cytological markers published till 31 December 2009 available in either Pubmed or Embase to detect or exclude the presence of malignant mesothelioma. Studies had to include ≥10 malignant mesothelioma patients and allow for construction of a two-by-two table. We assessed the study quality with use of the QUADAS criteria. In total, 82 studies fulfilled the criteria of this systemic review. The general quality of the incorporated studies was poor. Common methodological flaws included use of a non-representative patient sample and having an unknown delay between the index test and diagnosis. Furthermore, 88% of the studies were hampered by partial verification bias due to an inappropriate study design. The most frequently studied immunohistochemical markers for cytological analysis were EMA, Ber-Ep4, CEA, and calretinin. The most frequently investigated serum marker was SMRP. The markers CEA, Ber-EP4 and calretinin were most valuable in discriminating malignant mesothelioma from other malignant diseases. The markers EMA and SMRP were most valuable in discriminating malignant mesothelioma from non-malignant diseases. No marker performed well in discriminating between malignant mesothelioma and both malignant and non-malignant diseases. Moreover, our results showed considerable unexplained study heterogeneity, indicating that the evidence to properly assess the value of non-invasive marker tests in the diagnosis of malignant mesothelioma is currently limited.

Recently, more treatment options have become available for patients with malignant mesothelioma. Regularly updating survival estimates using population-based studies is therefore valuable. In chapter 4 we assessed the overall (baseline) survival as well as related prognostic variables in a large cohort of 1,353 patients with a confirmed diagnosis of malignant mesothelioma between 2005-2008. Additionally, the predictive accuracy of combined prognostic factors for survival was assessed. Of the patients about 50% were 70 years or older at time of diagnosis. The median latency time since start of asbestos exposure was 49 years. One year after diagnosis 47% of the patients were alive, 20% after two years and 15% after three years. Prognostic variables independently associated with worse survival were: older age (HR=1.04 per year 95%CI [1.03-1.06]), sarcomatoid subtype

(HR=2.45 95%CI [2.06-2.90]), and non-pleural localization (HR=1.67 95%CI [1.26-2.22]). The combined discriminative ability of these variables (c-index) was 0.66 (95%CI [0.64-0.68]). Overall, our results indicated that survival in patients with malignant mesothelioma is still limited, which is consistent with other population based studies.

Methodological considerations
Based on the experience of our diagnostic systematic review of markers for the detection of malignant mesothelioma, we discussed in chapter 7 various issues arising from heterogeneity in diagnostic test accuracy studies. We explained the meaning and relevance of the between-study variances and covariance parameters in the recently propagated bivariate model for diagnostic meta-analyses and described different ways of presenting study heterogeneity. Moreover, we discuss some (alternative) options when meta-analyzing continuous diagnostic tests in which different cut-offs have been applied (such as mesothelioma markers).

The *second part* of this thesis relates to the asbestos-related lung cancer risk in the general population. Exposure to asbestos is known to increase the risk of lung cancer but the exact relationship between exposure and risk is unknown. Most previous lung cancer studies focused on individuals heavily exposed to asbestos. Therefore, the existing estimated lung cancer risks per unit of exposure are mainly based on, and applicable to, high exposure levels. However, in the Western world handling of asbestos declined gradually after the 1970s and dropped severely in the 1990s due to directives on protecting workers exposed to asbestos. As a result, it is unlikely that individuals are currently exposed to the high levels previously studied. Predicting the impact of low exposures in the general population by estimating excess risk accurately at relatively low exposures is therefore becoming ever more relevant. In chapter 5, we provided new evidence by fitting flexible meta-regression models, a new method to assess the risk at low cumulative asbestos exposure. We fitted linear and non-linear meta-regression models to risk estimates of 104 exposure categories extracted from 19 asbestos lung cancer studies. Associated relative risks (RRs) were calculated for several low cumulative asbestos exposure levels. Based on an approach using natural splines the relative lung cancer risk for cumulative exposure levels of 4 f-y/ml, and 40 f-y/ml was estimated to range from 1.013 to 1.027, and from 1.13 to 1.30, respectively. After stratification by fibre type a non-significant 3 to 4-fold difference in RRs between chrysotile and amphibole fibres was found for exposures below 40 f-y/ml. Fibre type-specific risk estimates were strongly influenced by a few studies. In conclusion, our regression method indicated that at lower asbestos exposure levels the increase in RR of lung cancer due to asbestos exposure may be larger than expected from previous meta-analyses. Observed potency differences between different fibre types were lower than the generally held consensus.

To estimate the (expected future) number of lung cancers due to asbestos exposure, various methods have been applied. In chapter 6 we presented three modeling methods, applied to

the Dutch population, and discussed their evidence requirements, (dis)advantages, and the similarity of their results. The first method was relatively simple and required little evidence, estimating asbestos-related lung cancer cases directly from observed and predicted malignant mesothelioma cases in an Age-Period-Cohort (APC) analysis. The second method was slightly more complex also requiring evidence on the fraction of lung cancer cases attributable to asbestos exposure. The third method was the most comprehensive, requiring actual exposure information to perform a life table analysis on all individuals in the Dutch population. It is clear that the different modeling methods have different requirements in terms of both time to construct the model and evidence to populate the model. In our analysis of the number of future asbestos-related lung cancer cases in The Netherlands we found that the three methods produced different estimates: first method (based on the estimated number of mesotheliomas), 21,500-27,200 cases, second method (based on the population attributable risk) 12,150 cases, third method (life table analyses) 6,300 cases. The preferred method(s) for estimating asbestos-related lung cancer cases necessarily depends on the information and accuracy of the information that is available as model input. We show that using three different methods results in different absolute estimates varying by a factor of ~4. As such the exact impact of asbestos exposure on the lung cancer burden remains uncertain.

Samenvatting

Nederlandse samenvatting

Al vanaf 1930 is bekend dat asbest een levensgevaarlijke stof is. Toch duurde het nog vele decennia totdat het gebruik van asbest verboden werd in de meeste westerse landen. In niet-westerse landen neemt het gebruik van asbest momenteel zelfs toe.

Vandaag de dag is de impact van het gebruik van asbest in de westerse wereld nog goed merkbaar. Asbest is te vinden in vele gebouwen en blootstelling is mogelijk bij sloop en renovatie van deze gebouwen. Bovendien zijn de gezondheidseffecten pas vaak op langer termijn zichtbaar doordat asbest vele jaren na blootstelling tot verschillende soorten kankers kan leiden. De meest voorkomende asbest-gerelateerde kankers zijn maligne mesothelioom en longkanker. Verwacht wordt dat het aantal patiënten met maligne mesothelioom de komende jaren nog zal toenemen.

Het eerste deel van dit proefschrift beschrijft diagnostische en prognostische aspecten van maligne mesothelioom. Hoewel een definitieve diagnose van maligne mesothelioom belangrijk is voor de medische behandeling is het ook vereist om in aanmerking te komen voor een financiële vergoeding. In Nederland kunnen patiënten met ogenschijnlijk maligne mesothelioom zich aanmelden bij het Instituut voor Asbestslachtoffers voor een financiële compensatie. Voorafgaand aan deze compensatie dient de diagnose bevestigd te worden door een nationaal mesotheliomen panel van pathologen op grond van een cytologische of histologische analyse. Wanneer de diagnose niet pathologisch bevestigd kan worden, wordt op grond van beschikbare klinische gegevens en röntgenonderzoek een diagnose gesteld door een panel van klinische specialisten. Indien de diagnose is bevestigd aan de hand van pathologisch materiaal of klinische gegevens, komen patiënten in Nederland in aanmerking voor een vergoeding. Hoofdstuk 2 beschrijft het diagnostisch proces van patiënten die zich aanmelden bij het Instituut voor Asbestslachtoffers. Hierin wordt nagegaan hoe vaak een pathologische en klinische diagnose vastgesteld kan worden en welke aspecten van belang zijn bij een bevestiging van de diagnose. Bij 97 van de 1,498 patiënten (6%) die zich bij het Instituut tussen 2005 en 2008 hadden aangemeld kon geen pathologische diagnose worden vastgesteld. Redenen hiervoor waren dat de diagnose onzeker was door een differentiaaldiagnose (N=54) of dat het pathologisch materiaal onvoldoende was (N=22) of helemaal niet beschikbaar (N=21). Patiënten waarvan biopsie gegevens beschikbaar waren, kregen vaker een pathologische diagnose dan patiënten waarvan alleen cytologische gegevens beschikbaar waren. Van de patiënten waarvan de diagnose op grond van de beschikbare klinische gegevens moest worden vastgesteld, werd 80% beoordeeld als een diagnose maligne mesothelioom. Hierbij was geen enkel klinisch gegeven sterk geassocieerd met de diagnose maligne mesothelioom. Wanneer patiënten alleen een financiële vergoeding kunnen krijgen op grond van een pathologische diagnose zal een kleine groep patiënten niet de mogelijkheid hebben om compensatie te krijgen bij leven. Wanneer pathologisch materiaal onvoldoende is, zou een diagnose op grond van klinische gegevens

een goed alternatief kunnen zijn, vooral wanneer een re-biopsie geen optie meer is. Het vaststellen van een diagnose op basis van klinische gegevens is een multivariabel proces en dient dan ook zorgvuldig door meerdere specialisten te worden gedaan. Om de kans op een bevestigde pathologische diagnose te verhogen is het aan te bevelen om, indien mogelijk, altijd biopsie materiaal af te nemen bij patiënten. Zodoende is het belangrijk dat patiënten snel worden doorverwezen voor diagnostische procedures.

Om het diagnostisch proces te versnellen zijn talrijke markers geëvalueerd teneinde de niet-invasieve diagnostische work-up bij patiënten met maligne mesothelioom te vergemakkelijken. Hoofdstuk 3 beschrijft een systematische review naar de diagnostische waarde van niet-invasieve marker testen voor het vaststellen of uitsluiten van de diagnose maligne mesothelioom. Studies gepubliceerd in PubMed of Embase tot en met 31 December 2009 werden geselecteerd indien ze de diagnostische waarde van serum of cytologische markers bij maligne mesothelioom hadden onderzocht. De studies moesten minstens 10 maligne mesothelioom patiënten omvatten en voldoende resultaten hebben om een 2 bij 2 tabel te maken. De kwaliteit van de studies werd beoordeeld met de QUADAS criteria. In totaal leverde de systematische review 82 studies op. De kwaliteit van de studies was in het algemeen laag. Voorkomende tekortkomingen waren een niet-representatieve steekproef en onduidelijkheid over de tijd tussen de marker test en de uiteindelijke diagnose. Bovendien was in 88% van de studies sprake van partiële verificatie bias door een onjuist studie design. Voor cytologische analyse waren de immunohistochemische markers EMA, Ber-EP4, CEA en calretinine het meest onderzocht. SMRP was de meest onderzochte serum marker. De markers CEA, Ber-EP4 en calretinine waren het meest waardevol om maligne mesothelioom te onderscheiden van andere kwaadaardige tumors. De markers EMA en SMRP waren het meest waardevol om maligne mesothelioom van goedaardige aandoeningen te onderscheiden. Geen enkele test was voldoende klinisch waardevol om als zelfstandig diagnostisch instrument te kunnen worden gebruikt. Verder was er grote (onverklaarde) heterogeniteit tussen studies, waardoor de diagnostische waarde van serum en cytologische markers in de klinische praktijk lastig was te kwantificeren.

Recentelijk zijn er meer mogelijkheden voor de behandeling van maligne mesothelioom. Het is dan ook zinvol om de levensverwachting van patiënten met maligne mesothelioom regelmatig opnieuw te schatten aan de hand van populatie gegevens. In hoofdstuk 4 beschrijven we de prognose en prognostische factoren in een groot cohort van 1,353 patiënten met maligne mesothelioom. Bij deze patiënten werd de diagnose maligne mesothelioom in 2005-2008 vastgesteld. Op het moment van diagnose was ongeveer de helft van de patiënten 70 jaar of ouder. De mediane latentietijd tussen diagnose en eerste blootstelling aan asbest bedroeg 49 jaar. Na diagnose overleefde 47% van de patiënten het eerste jaar. Slechts 20% van de patiënten overleefde meer dan 2 jaar, en 3 jaar na de diagnose was nog 15% van de patiënten in leven. Factoren die onafhankelijk geassocieerd waren met een slechtere prognose, waren: oudere leeftijd (HR=1.04 per jaar

95%BI [1.03-1.06]), sarcomatoide subtype (HR=2.45 95%BI [2.06-2.90]) en een peritoneale ligging (HR=1.67 95%BI [1.26-2.22]). De discriminatie (de c-index) van het model waarin deze variabelen gecombineerd werden, was 0.66 (95% BI [0.64-0.68]). Concluderend, de overleving bij maligne mesothelioom is nog steeds kort, hetgeen overeenkomt met andere populatie studies.

Methodologische aspecten

Op basis van onze ervaring met het diagnostisch systematisch review beschrijven we in hoofdstuk 7 verschillende aspecten van heterogeniteit in diagnostische studies. We lichten de betekenis en relevantie van de tussen-studie varianties en covariantie parameters toe van het vrij nieuwe bivariaat model en beschrijven de verschillende mogelijkheden om heterogeniteit tussen diagnostische studies te presenteren. Tevens beschrijven we hoe continue testen waarbij verschillende afkappunten zijn gehanteerd (zoals markers voor maligne mesothelioom), kunnen worden geanalyseerd.

Het tweede deel van dit proefschrift beschrijft het risico van asbest-gerelateerde longkanker in de algemene bevolking. Het is algemeen bekend dat asbest geassocieerd is met longkanker, echter de exacte relatie tussen asbest en longkanker is onbekend. De meeste studies uitgevoerd tot nu toe hebben het longkankerrisico onderzocht bij personen met een hoge blootstelling aan asbest. Bestaande schattingen van het longkankerrisico per vezeljaar zijn daarom gebaseerd en toepasbaar op personen met hoge blootstellingen. Echter, het gebruik van asbest in de westerse wereld daalde na 1970 en werd verder teruggebracht in de jaren 90 door maatregelingen om werknemers te beschermen tegen asbestblootstelling. Het is daardoor niet aannemelijk dat de huidige bevolking in de westerse wereld een cumulatieve blootstelling heeft die zo hoog is als in de vroegere studies. Om de impact van asbest op longkanker in de huidige bevolking na te gaan is het dus noodzakelijk om het risico op longkanker bij lage blootstellingen goed te kunnen schatten. In hoofdstuk 5 beschrijven we nieuwe resultaten aan de hand van een meer flexibele meta-regressie methode. Met deze methode kan het longkankerrisico bij lage asbestblootstellingen geschat worden. Lineaire en non-lineaire modellen zijn gefit op 104 datapunten (risicoschattingen). Deze punten werden uit 19 studies geëxtraheerd en gaven het aggregeerde longkankerrisico bij een bepaalde blootstelling categorie weer. Het relatief risico (RR) op longkanker, zoals geschat volgens deze modellen, werd berekend voor verschillende lage cumulatieve blootstellingen aan asbest. Op basis van een spline model werd het relatief risico voor 4 en 40 vezeljaren geschat tussen 1.013 en 1.027 en 1.13 en 1.30, respectievelijk. Na stratificatie naar vezeltype werd een 3 tot 4 keer niet-significant hoger relatief risico geschat bij blauw asbest dan bij wit asbest voor blootstellingen onder de 40 vezeljaren. De geschatte verschillen in het risico tussen vezeltypen waren sterk afhankelijk van een paar studies. De gebruikte regressiemethode toont aan dat het longkankerrisico bij lage blootstellingen hoger is dan gerapporteerd in eerdere meta-analyses. Verschillen tussen vezeltypes lijken kleiner te zijn dan vaak wordt aangenomen.

In hoofdstuk 6 worden verschillende modellen gepresenteerd die zijn toegepast op de Nederlandse situatie om een schatting te kunnen geven van het aantal (toekomstige) asbest-gerelateerde longkanker gevallen. We beschrijven de benodigde invoer, de voor- en nadelen en hoe de resultaten van de drie modellen overeenkomen. In het eerste model, een relatief simpel model, werd het aantal toekomstige longkanker gevallen direct afgeleid van het aantal geschatte maligne mesothelioom gevallen in een leeftijd-periode-cohort analyse. In het tweede model was informatie nodig over de fractie longkanker gevallen dat te wijten is aan asbestblootstelling (het populatie attributief risico). Het derde model was het meest uitgebreide model, waarin gegevens over blootstellingen in de algemene bevolking vereist waren om een overlevingstabel analyse uit te voeren op de Nederlandse bevolking. Deze drie methoden verschilden duidelijk van elkaar in termen van benodigde tijd en data om het model op te zetten. In de analyse van het aantal toekomstige asbest-gerelateerde longkanker gevallen liepen de schattingen van de modellen ook beduidend uiteen: het eerste model (op basis van het aantal verwachte mesotheliomen) schatte 21,500 - 27,200 gevallen, het tweede model (het populatie attributief risico) schatte 12,150 gevallen, en het derde model (de overlevingstabel analyse) schatte 6,300 gevallen. In de praktijk zal het te prefereren model met name afhangen van de data die beschikbaar is en de accuraatheid daarvan. In deze studie laten we zien dat wanneer er verschillende modellen worden gebruikt om het aantal asbest-gerelateerde longkanker gevallen te schatten deze schattingen kunnen variëren met een factor ~4. Hierdoor is de impact van asbestblootstelling op de omvang van longkanker moeilijk te kwantificeren.

Dankwoord

Dankwoord

Promoveren is net als een verre reis. Tijdens je reis maak je van alles mee: de ene dag zit je op een tropisch eiland, de andere dag bevind je je in een bus over een hobbelige weg. Om je bestemming te bereiken, moet je soms obstakels overwinnen. Eenmaal goed aangekomen weet je waarvoor je het doet. Met dit woord van dank sluit ik een enerverende reis af en begin ik aan een nieuwe episode in mijn leven.

Graag wil ik enkele personen in het bijzonder bedanken, die mij hebben ondersteund bij dit proefschrift. Allereerst mijn promotoren en co-promotoren. Zonder hun deskundige begeleiding was ik niet ver gekomen.

Dr. ir. Hendrik Koffijberg, beste Erik, het zal je niet verbazen dat ik jou als eerste noem. Jij hebt mij intensief geholpen bij het voltooien van dit proefschrift. Ik vond het ontzettend prettig en leerzaam om jou als co-promotor te hebben. Je wist me altijd me weer te motiveren en een stuk tekst na jouw commentaar leverde elke keer een nog betere versie op, speciale dank daarvoor.

Dr. ir. Roel Vermeulen, beste Roel, leuk dat jij mijn tweede co-promotor bent geworden. Jouw inbreng en snelle reacties zorgden voor een enorme boost in het proces. Ik bewonder je kennis en enthousiasme voor het onderzoek. Ik wil je dan ook hartelijk bedanken voor je ondersteuning.

Prof. dr. mr. dr. Bas de Mol, beste Bas, ik dank je voor de impulsen, die je aan dit onderzoek hebt gegeven. Jouw goede aanvullingen en sturing hebben mede voor dit eindresultaat gezorgd. Met plezier denk ik terug aan de boeiende gesprekken en jouw tomeloze energie.

Last but not least, Prof. dr. Karel Moons. Beste Carl, volgens mij kan geen enkele promovendus zich een betere promotor wensen. Je hebt een belangrijke rol gespeeld in de sturing van het project. Goed waren de gesprekken in de tijd dat het tegenzat. Alle lof voor jouw volle vertrouwen, jouw geduld wanneer ik een beer op de weg zag die er niet was, en jouw heldere en kritische inzichten.

Mijn medeauteurs, Sjaak Burgers, Prof. Paul Baas, Prof. Marc van de Vijver, Eva Schaake, Lützen Portengen, Virissa Lenters, Prof. Dick Heederik, Hans Reitsma en Peter Zuithoff voor hun belangrijke inbreng in dit proefschrift. Sjaak, wat was het ontzettend prettig dat ik jou altijd kon bellen en mailen voor vragen over de praktijk. Het dagje meelopen heb ik als bijzonder ervaren. Jouw gevoel voor humor maakte onze samenwerking extra plezierig. Hans, ik waardeer het enorm dat jij betrokken bent geweest bij mijn laatste artikel. Jouw inspiratie werkt aanstekelijk. Lützen, enorme dank voor al jouw uitleg en bijdrage in een groot deel van de statistische analyses.

Natuurlijk mijn collega's op het JC, in het bijzonder Maartje (niet voor niets mijn paranimf) en Mariëtte, voor de niet altijd onderzoeksgerelateerde maar wel gezellige onderonsjes.

Verder wil ik alle stuurgroepleden bedanken voor hun betrokkenheid bij het onderzoek en Jan Tempelman voor zijn deskundigheid over asbestblootstelling.

Ik sluit af met mijn familie. Lieve heit en mem, wat ontzettend fijn dat jullie er altijd voor mij zijn. Jullie stimulans en onvoorwaardelijke liefde heeft ervoor gezorgd dat ik ben waar ik nu ben. Mijn zussen, Akke, Jacomien en Hendrikje (mijn paranimf). Ik kan me geen leukere zussen voorstellen. Wat is het heerlijk om leuke dingen met jullie te ondernemen om even de zinnen te verzetten.

Lieve Ramon, Jij hebt in alle omstandigheden mij ondersteund het proefschrift af te maken. Ik ben jou vele dank (en tijd) verschuldigd. Gelukkig is het nu af, zodat we eindelijk tijd hebben: tijd voor elkaar. Ik heb er zin in!

Lieve Luna, wat is het een voorrecht om te zien hoe jij op jouw manier op ontdekkingsreis bent in deze grote wereld.

Nu is het boek af, op naar een volgend avontuur!

Sjoukje

Curriculum Vitae

Curriculum Vitae

Sjoukje van der Bij was born on December 24th, 1978 in Arnhem, the Netherlands. She was raised in Akkrum, Friesland and in 1997, she completed her secondary education at the Bornego College in Heerenveen. The following five years she studied Biomedical Sciences, major Epidemiology, at the University of Nijmegen. In 2004 she started her working career as researcher at the Tympaan Institute, an institute in The Hague for policy development in the province South Holland. She switched to pharmaco-epidemiologic research using large data base registries at the PHARMO Institute Outcomes Research, Utrecht in 2005. In 2008 she started the work described in this thesis at the Julius Center for Health Sciences and primary Care, University Medical Center Utrecht, The Netherlands under the supervision of Prof. dr. K.G.M. Moons, Prof. dr. mr. dr. B.A.J.M. de Mol, Dr. ir. Koffijberg and Dr. ir. R.C.M. Vermeulen. As of July 2012, she is working as a researcher at the NIVEL (Netherlands institute for health services research) where she is involved in research on the quality of data registration in electronic patients records in the general health care.